# Bayes'sche Entscheidungstheorie

Inaugural-Dissertation
zur Erlangung der Würde eines Doctor rerum politicarum
an der Rechts- und Wirtschaftswissenschaftlichen
Fakultät der Universität Bern

vorgelegt von

Benedikt Weibel
von Rapperswil (BE)

1977

Verlag Paul Haupt Bern 1978

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548908_0008

# Bayes'sche Entscheidungstheorie

Inaugural-Dissertation
zur Erlangung der Würde eines Doctor rerum politicarum
an der Rechts- und Wirtschaftswissenschaftlichen
Fakultät der Universität Bern

vorgelegt von

Benedikt Weibel
von Rapperswil (BE)

UNIV. LIBRARY BERN

Verlag Paul Haupt Bern 1978

Die Fakultät hat diese Arbeit am 15. Dezember 1977 auf Antrag von Herrn Prof. Dr. W. Müller als Dissertation angenommen, ohne damit zu den darin angesprochenen Auffassungen Stellung nehmen zu wollen.

UNIV.
LEIDEN
BIBL.

Diese Arbeit erscheint unter dem gleichen Titel als Band 7 der Schriftenreihe Prüfen und Entscheiden im Verlag Paul Haupt Bern und Stuttgart

ISBN 3-258-02793-5

Alle Rechte vorbehalten
Copyright © 1978 by Paul Haupt Berne

VORWORT

Meinem Doktorvater, Herrn Prof. Dr. W. Müller, danke ich für manche wertvolle Anregung und dafür, dass er diese Arbeit in seiner Schriftenreihe aufgenommen hat. Zu ganz besonderem Dank bin ich aber auch Herrn Dr. Tran-Ngoc-An verpflichtet, der mir bei den mathematischen Problemen beigestanden hat und der mit grosser Geduld die Arbeit in formeller Hinsicht zu wiederholten Malen überprüft hat.

INHALTSVERZEICHNIS


| | Seite |
|---|---|
| Abbildungsverzeichnis | V |
| Tabellenverzeichnis | VII |
| EINLEITUNG | 1 |
| TEIL I BAYES'SCHE ENTSCHEIDUNGSTHEORIE ALS TEILGEBIET DER ENTSCHEIDUNGSTHEORIE | 3 |
| 1. Gegenstand und Grundproblem der Entscheidungstheorie | 3 |
| 2. Das allgemeine Entscheidungsmodell | 5 |
| 3. Klassifikation von Entscheidungsmodellen | 9 |
| 4. Einführung in das Bayes'sche Modell des Entscheidens | 10 |
| 5. Prämissen | 20 |
| TEIL II EINE TYPOLOGIE VON BAYES-ENTSCHEIDUNGEN | 22 |
| 1. Typologische Merkmale | 23 |
| 1.1. Die Informationsstruktur | 23 |
| 1.2. Das stochastische Modell | 23 |
| 1.21. Das Binomial-Modell | 24 |
| 1.22. Das Hypergeometrische Modell | 25 |
| 1.23. Das Poisson-Modell | 26 |
| 1.24. Das Normal-Modell | 27 |
| 1.3. Die Apriori-Verteilung | 30 |
| Exkurs: Deutungen des Wahrscheinlichkeitsbegriffs | 31 |
| a) Nominaldefinitionen der Wahrscheinlichkeit | 32 |
| b) Die drei klassischen Wahrscheinlichkeitskonzepte | 34 |
| ba) Die logische Wahrscheinlichkeit | 34 |
| bb) Die objektive Wahrscheinlichkeit | 35 |
| bc) Die subjektive Wahrscheinlichkeit | 36 |

c) Monistische und dualistische Auffassungen 39

d) Kritik der einzelnen Konzeptionen 40

e) Wahrscheinlichkeit in der Entscheidungstheorie 43

1.31. Interpretation der Apriori-Verteilung 45

1.32. Die Apriori-Verteilung in der Bayes'schen Entscheidungstheorie 46

1.33. Ausprägungen der Apriori-Verteilung 48

1.331. Diskrete Apriori-Verteilung 48

1.332. Stetige Apriori-Verteilung 48

1.332.1. Treppenverteilung 49

1.332.2. Konjugierte Familien von Verteilungen 50

1.332.21. Binomial-Modell: Betaverteilung 52

1.332.22. Poisson-Modell: Gammaverteilung 54

1.332.23. Normal-Modell 55

a) $\sigma$ bekannt, $\mu$ unbekannt: Normalverteilung 57

b) $\mu$ bekannt, $\sigma$ unbekannt: Gamma-2-Verteilung 58

1.4. Die Art der Konsequenzenfunktion 59

2. Kombination der typologischen Merkmale 60

TEIL III DIE ENTSCHEIDUNG 63

1. Die Entscheidung ohne zusätzliche Informationsaktivität 63

2. Die Entscheidung mit zusätzlichem Informationsumfang: gegebener Informationsumfang 71

2.1. Zwei Methoden der Analyse des Problems 72

2.11. Die Normalform der Analyse 73

2.12. Die extensive Form der Analyse 78

2.13. Vergleich der beiden Methoden 79

2.2. Die Entscheidung bei diskreter Apriori-Verteilung 81

2.21. Binomial-Modell 81

2.22. Hypergeometrisches Modell 88

2.23. Poisson-Modell 92

2.24. Normal-Modell 97

2.25. Ueber den Realitätsgehalt von Entscheidungsmodellen: Kritik der bisher verwendeten Modelle und Vorschläge für deren Ausbau 111

2.3. Die Entscheidung bei stetiger Apriori-Verteilung 116

2.31. Treppenverteilung 116

2.311. Das Testen von Hypothesen in der klassischen Statistik 116

2.312. Eine "Bayesianische Anwendung" des Hypothesen-Tests 120

2.32. Konjugierte Familien von Verteilungen 124

3. Die Entscheidung mit zusätzlicher Informationsaktivität: variabler Informationsumfang 131

3.1. Der Wert zusätzlicher Information 132

3.2. Die optimale Strategie 134

3.3. Die Entscheidung bei diskreter Apriori-Verteilung 135

3.31. Binomial-Modell 135

3.32. Hypergeometrisches Modell 139

3.33 Poisson-Modell 144

3.34. Normal-Modell 149

3.4. Die Entscheidung bei stetiger Apriori-Verteilung 152

3.41. Treppenverteilung 152

3.42. Konjugierte Familien von Verteilungen 152

3.421. Allgemeine Theorie 154

3.421.1. Die Ermittlung des EVI 154

3.421.2. Die Ermittlung des EIS 155

3.421.3. Die Ermittlung der optimalen Strategie 155

3.422. Normal-Modell 156

3.422.1. Die Ermittlung des EVI 156

3.422.2. Die Ermittlung des EIS 157

3.423. Binomial-Modell 159

3.423.1. Die Ermittlung des EVI 159

3.423.2. Die Ermittlung des EIS 161

3.424. Poisson-Modell 164

3.424.1. Die Ermittlung des EVI 164

3.424.2. Die Ermittlung des EIS 166

3.425. Abschliessende Bemerkungen zum Konzept der konjugierten Familien von Verteilungen 167

4. Die Entscheidung mit zusätzlicher Informationsaktivität: variabler Informationsumfang; sequentieller Informationsentscheid 168

ANHANG 170

1. Herleitung von k* für das Normal-Modell 170

2. Herleitung von k* für das Normal-Modell als Approximation des Binomial-Modells 172

Literaturverzeichnis 176

ABBILDUNGSVERZEICHNIS


| Abbildung | | Seite |
|---|---|---|
| 1 | Darstellung des allgemeinen Entscheidungsmodells mit einem Entscheidungsbaum | 8 |
| 2 | Das Bayes'sche Modell des Entscheidens | 12 |
| 3 | Entscheidungsbaum der $e_o$-Strategie | 18 |
| 4 | Diskrete Apriori-Verteilung | 48 |
| 5 | Stetige Apriori-Verteilung | 49 |
| 6 | Treppenverteilung über zwei Intervallen | 50 |
| 7 | Konsequenzenfunktionen und Break-even-Punkt eines 2-Aktionen-Problems | 66 |
| 8 | Opportunitätskostenfunktionen eines 2-Aktionen-Problems | 67 |
| 9 | Treppenverteilung mit den Intervallen $[0,s_b]$ und $[s_b,1]$ | 69 |
| 10 | Opportunitätskostenfuntion einer Aktion $a_1$ | 71 |
| 11 | Entscheidungsbaum für die Entscheidungsregel $d_1$ und n=2 bei Normalform der Analyse | 76 |
| 12 | Entscheidungsbaum für die extensive Form der Analyse | 78 |
| 13 | Entscheidungsbaum für das Qualitätskontroll-Beispiel; extensive Form der Analyse | 86 |
| 14 | Entscheidungsbaum für das Autoverkäufer-Beispiel - exemplarisch für $d_4$ | 94 |
| 15 | Entscheidungsbaum für das Verpackungs-Beispiel | 100 |
| 16 | α- und β-Fehler beim Verpackungs-Beispiel | 102 |

| | | |
|---|---|---|
| 17 | Entscheidungsbaum für den Fall der PAN ATLANTIC Airways | 108 |
| 18 | $\alpha$- und $\beta$-Fehler beim Fall der PAN ATLANTIC Airways | 109 |
| 19 | Irrtumscharakteristiken des Verpakkungsproblems (n=100) | 119 |
| 20 | Entscheidungsbaum für den Fall stetige Apriori-Verteilung/stetige Konsequenzenfunktion | 126 |
| 21 | Konjugierte Apriori- und Aposteriori-Verteilungen | 128 |
| 22 | EVI, EIS(n), c(n), ENGS(n) und Eo(n) für das Qualitätskontroll-Beispiel | 138 |
| 23 | Entscheidungsbaum für das Flugzeugbeispiel | 140 |

TABELLENVERZEICHNIS


| Tabelle | | Seite |
|---|---|---|
| 1 | Darstellung des allgemeinen Entscheidungsmodells in Tabellenform | 6 |
| 2 | Formen von Betaverteilungen | 54 |
| 3 | Die typologischen Merkmale und ihre Ausprägungen | 61 |
| 4 | Konsequenzentabelle eines 2-Aktionen-Problems | 72 |
| 5 | Mögliche Entscheidungsregeln bei einem 2-Aktionen-Problem und einer Stichprobe $e_2$ | 73 |
| 6 | Sinnvolle Entscheidungsregeln bei einem 2-Aktionen-Problem und einer Stichprobe $e_2$ | 75 |
| 7 | Entscheidungsregeln und ihre bedingten Konsequenzenerwartungswerte | 77 |
| 8 | Konsequenzentabelle für das Qualitätskontroll-Beispiel | 81 |
| 9 | Opportunitätskostentabelle für das Qualitätskontroll-Beispiel | 81 |
| 10 | Entscheidungsregeln und ihre erwarteten Opportunitätskosten für das Qualitätskontroll-Beispiel | 84 |
| 11 | Entscheidungsregeln und ihre Opportunitätskosten-Erwartungswerte für das Qualitätskontroll-Beispiel mit einer Stichprobe ohne Zurücklegen | 90 |
| 12 | Opportunitätskostentabelle für das Autoverkäufer-Beispiel | 93 |
| 13 | Entscheidungsregeln und ihre erwarteten Opportunitätskosten für das Autoverkäufer-Beispiel | 96 |
| 14 | Konsequenzentabelle für das Verpackungs-Beispiel | 98 |

| | | |
|---|---|---|
| 15 | Opportunitätskostentabelle für das Verpackungs-Beispiel | 99 |
| 16 | Entscheidungsregeln und ihre erwarteten Opportunitätskosten für das Verpackungs-Beispiel | 103 |
| 17 | Opportunitätskostentabelle für das Problem der PAN ATLANTIC Airways | 107 |
| 18 | Opportunitätskostentabelle mit als zusammengesetzten Hypothesen formulierten Umweltzuständen | 116 |
| 19 | Opportunitätskostentabelle eines Qualitätskontroll-Beispiels | 135 |
| 20 | Die ENGS(n) für das Qualitätskontroll-Beispiel | 137 |
| 21 | Konsequenzentabelle für das Flugzeugbeispiel | 139 |
| 22 | Rechenschema für die Ermittlung des Konsequenzenerwartungswertes der $e_{10}$-Strategie für das Flugzeugbeispiel | 142 |
| 23 | Die ENGS(n) für das Flugzeugbeispiel | 144 |
| 24 | Konsequenzentabelle für das Maschinenbeispiel | 145 |
| 25 | Opportunitätskostentabelle für das Maschinenbeispiel | 147 |
| 26 | Die ENGS(t) für das Maschinenbeispiel | 149 |
| 27 | Die ENGS(n) für das Verpackungsbeispiel | 152 |

## Einleitung

Der Gegenstand der vorliegenden Arbeit, die Bayes'sche Entscheidungstheorie, ist im Ueberlappungsbereich zweier Disziplinen anzusiedeln: Einmal, wie aus ihrem Namen ersichtlich, im Bereiche der allgemeinen Entscheidungstheorie und zum zweiten in demjenigen der Statistik, weshalb etwa auch von "Bayes'scher Statistik" die Rede ist. Die Erweiterung des allgemeinen Entscheidungsmodells[1] um die Dimension der zusätzlichen Information über relevante Umweltzustände erfordert statistische Methoden, um zu einer Entscheidung zu gelangen. Anderseits hat ein Grundprinzip entscheidungstheoretischen Denkens, die Berücksichtigung der Folgen bzw. Konsequenzen von Handlungen bei einer Entscheidung, der Statistik einen neuen Ansatz und eine neue Schule, eben diejenige der Bayesianer, gebracht.

Im ersten Teil der Arbeit wird die Bayes'sche Entscheidungstheorie in den Rahmen der allgemeinen Entscheidungstheorie gestellt. Damit soll eine Situierung der in dieser Untersuchung behandelten Probleme ermöglicht und eine für den zweiten Teil benötigte Terminologie bereitgestellt werden. Der Abschnitt "Einführung in das Bayes'sche Modell des Entscheidens" dient insbesondere diesem terminologischen Zweck und erläutert die erforderlichen Begriffe anhand eines einfachen Modells einer Bayes-Entscheidung[2]. Es erfüllt dieser Abschnitt also nicht den Anspruch einer umfassenden Darstellung des Grundproblems der Bayes'schen Entscheidungstheorie, dies ist erst nach der im zweiten Teil entwickelten Typologie möglich.

In diesem zweiten Teil geht es darum, das grob skizzierte Problem so zu strukturieren, dass Problemtypen entstehen,

[1] Vgl. S. 5 ff.

[2] Zum Begriff der Bayes-Entscheidung s.S. 18

die spezifischen Lösungsmethoden zugänglich sind. Es werden zunächst strukturierende oder typologische Merkmale aufgezählt und ausführlich beschrieben. Kombinationen dieser Merkmale ergeben anschliessend die verschiedenen Problemtypen.

Der dritte und letzte Teil ist schliesslich der Darstellung der Lösungsmethoden, also der eigentlichen Entscheidung, gewidmet. Der statistische Ansatz der Bayesianer wird in diesem Teil explizit gemacht und der herkömmlichen Betrachtung der Statistik gegenübergestellt.

# TEIL I BAYES'SCHE ENTSCHEIDUNGSTHEORIE ALS TEILGEBIET DER ENTSCHEIDUNGSTHEORIE

## 1. Gegenstand und Grundproblem der Entscheidungstheorie

Gegenstand der Entscheidungstheorie ist die Entscheidung, verstanden als ein bewusster "Wahlakt, der ein Handeln auslöst oder nicht auslöst und im ersteren Fall auch die Art und Weise des Handelns bestimmt"[1]. Nach der Formulierung des Grundproblems können zwei verschiedene Arten von Entscheidungstheorie unterschieden werden: Die deskriptive Entscheidungstheorie, die beobachtbare Wahlakte zu erklären versucht, also auf die Gewinnung beschreibender Aussagen ausgerichtet ist und die präskriptive Entscheidungstheorie, die Aussagen vorschreibender (normativer) Art aufstellt, also Empfehlungen abgibt, die dem Entscheidenden helfen sollen, die Wahlakte in einer für ihn möglichst günstigen Art zu treffen oder - um einen zentralen Begriff der präskriptiven Entscheidungstheorie zu verwenden - rationale Entscheide zu fällen. Das Rationalitätspostulat der präskriptiven Entscheidungstheorie bezieht sich nach der heute vorherrschenden Meinung ausschliesslich auf eine formale, methodische Rationalität, die keine Anforderungen an den Zielinhalt stellt, sondern nur die Form des Zielsystems und die Methoden seiner bestmöglichen Verwirklichung betrifft[2].

Bisweilen wird der präskriptiven und deskriptiven Entscheidungstheorie noch eine dritte Art, die Entscheidungslogik, beigefügt[3]. Diese befasst sich mit der Aufgabe, ausgehend von bestimmten Prämissen auf rein logisch-deduk-

[1] Müller, S. 2

[2] Vgl. Bamberg/Coenenberg, S. 3; Kühn, S. 6

[3] Gäfgen, S. 52 f.; Müller, S. 31

tivem Wege die rationale Entscheidung herzuleiten. Entscheidungslogik kann aber auch als Teilbereich der präskriptiven Entscheidungstheorie aufgefasst werden[1], denn ihr Ziel ist ohne Zweifel die Gewinnung normativer Aussagen. Den zweiten Bereich der präskriptiven Entscheidungstheorie könnte man dann "Entscheidungsheuristik" nennen. Ihr zentrales Element ist der Entscheidungsprozess und seine Zergliederung in einzelne Phasen[2]. Die dadurch gewonnenen Aussagen sind Angaben über Verfahrensschritte, "aus denen letztlich ein Handeln des Aktors resultiert, das als rational begründet qualifiziert werden kann"[3]. Der Terminus "Heuristik" soll zum Ausdruck bringen, dass dem Entscheidenden bei der Suche nach rationalen Entscheiden nichtdeduktive methodische Hilfe angeboten wird[4]. Dass es sich dabei nicht um eine Folge von logisch deduzierbaren Schritten handelt, verdeutlicht bereits die Tatsache der verschiedenen nebeneinander bestehenden Phasengliederungen in der Literatur.

Die Bayes'sche Entscheidungstheorie, mit der sich diese Arbeit befasst, ist der Entscheidungslogik zuzuordnen. Aus diesem Grunde befasst sich der folgende Abschnitt zunächst mit dem allgemeinen Modell der Entscheidungstheorie (eigentlich: der Entscheidungslogik), auch "Grundmodell" oder "Basismodell" des Entscheidens genannt[5].

---

[1] Bamberg/Coenenberg sehen sie sogar im Mittelpunkt der präskriptiven Entscheidungstheorie, Bamberg/Coenenberg, S. 2

[2] Vgl. etwa Kühn, S. 12 ff.; Müller, S. 17 ff; Szyperski/Winand, S. 6 ff.

[3] Müller, S. 15

[4] Es ist nicht die Intention des Verfassers, den überaus komplexen Begriff "Heuristik" mit dieser Kurzformel abschliessend zu definieren. Für die hier beabsichtigte Unterteilung der präskriptiven Entscheidungstheorie in Entscheidungslogik und Entscheidungsheuristik ist diese knappe Umschreibung m.E. aber ausreichend.

[5] Vgl. Bamberg/Coenenberg, S. 12; Bühlmann/Loeffel/Nievergelt, S. 1; Menges (2), S. 76

## 2. Das allgemeine Entscheidungsmodell

Das allgemeine Entscheidungsmodell besteht aus den folgenden konstitutiven Merkmalen[1]:

- Dem (oder den) Entscheidenden oder Aktor.
- Dem Ziel oder Zielsystem des Aktors.
- Den Handlungsmöglichkeiten, Alternativen oder Aktionen. Die Menge aller Aktionen $a_i$, $i=1,2,\ldots,I$ sei mit

  $A = \{a_1,\ldots,a_i,\ldots,a_I\}$

  bezeichnet.
- Den Zuständen der Umwelt oder unkontrollierbaren Situationsvariablen $s_j$, $j=1,2,\ldots,J$

  $S = \{s_1,\ldots,s_j,\ldots,s_J\}$.

  Darunter sind jene situationsrelevanten Gegebenheiten der Umwelt verstanden, die zusammen mit den Aktionen das Ergebnis einer Entscheidung beeinflussen.
- Den Wahrscheinlichkeiten

  $p(s_j), j=1,\ldots,J,$

  die den einzelnen Umweltzuständen zugeordnet sind.
- Den Handlungsfolgen, Konsequenzen oder Ergebnissen

  $k_{ij}$, $i=1,\ldots,I$; $j=1,\ldots,J$

  der Durchführung einer Aktion.
- Der Entscheidungsmaxime, die "besagt, wie aufgrund der Konsequenzenwerte gewählt werden muss"[2].

Das allgemeine Entscheidungsmodell wird üblicherweise

---

[1] Vgl. Bamberg/Coenenberg, S. 12 ff.; Menges (2), S. 82 ff.; Müller, S. 4 ff.

[2] Müller, S. 15

mittels Tabellen oder Entscheidungsbäumen dargestellt[1]. Die Tabelle heisst Konsequenzen- oder Ergebnistabelle und hat die folgende allgemeine Form:

| | $s_1$ . . . $s_j$ . . . $s_J$ |
|---|---|
| $a_1$ | $k_{11}$ . . $k_{1j}$ . . $k_{1J}$ |
| . | . . . |
| $a_i$ | $k_{i1}$ $k_{ij}$ . . $k_{iJ}$ |
| . | . . . |
| $a_I$ | $k_{I1}$ $k_{Ij}$ . . $k_{IJ}$ |
| | $p(s_1)$. . $p(s_j)$ . . $p(s_J)$ |

Tab. 1: Darstellung des allgemeinen Entscheidungsmodells in Tabellenform

Eine wichtige Eigenschaft von Konsequenzentabellen soll hier behandelt werden: Eine Konsequenzentabelle, die auf einer Ratio- oder Intervallskala gemessene Konsequenzen enthält[2], kann nach bestimmten Regeln so umgeformt werden, dass ihr Gehalt erhalten bleibt. Unter "Gehalt" wird dabei die Reihung der Aktionen in einer Präferenzordnung, die sich am Ziel bzw. Zielsystem orientiert und mit Hilfe einer Entscheidungsmaxime erstellt wird, verstanden. Tabellen, die in der betrachteten Art zusammenhängen, heis-

[1] Genau besehen wird durch die Tabelle bzw. den Entscheidungsbaum nur der Zusammenhang zwischen Aktionen, Umweltzuständen, Wahrscheinlichkeiten und Konsequenzen dargestellt. Ueber den Aktor, das Ziel und die Entscheidungsmaxime wird dabei nichts ausgesagt. Trotzdem soll im folgenden jeweils von der Darstellung des allgemeinen Entscheidungsmodells die Rede sein.

[2] Ueber die verschiedenen Messskalen und die Problematik der Messung vgl. Gäfgen, S. 144 ff.

sen "äquivalent"[1].

Es werden nun drei Regeln vorgestellt, nach denen äquivalente Tabellen generiert werden können, wenn die für die Bayes'sche Entscheidungstheorie zentrale Entscheidungsmaxime - die Erwartungswertmaxime - gegeben ist[2].

- Multiplikation aller Konsequenzenwerte mit derselben positiven Zahl.
- Addition einer bestimmten Zahl (positiv oder negativ) zu jedem Konsequenzenwert der Tabelle.
- Subtraktion des maximalen Konsequenzenwertes jeder Spalte von sämtlichen Konsequenzenwerten der entsprechenden Spalte.

Multipliziert man die nach der dritten Regel erhaltenen Werte mit -1, so nennt man die so entstandene Tabelle Opportunitätskostentabelle und ihre Elemente Opportunitätskosten $o_{ij}$, entgangene Gewinne oder Fehlentscheidungskosten.

Die Umformung einer Konsequenzentabelle kann u.U. den mit der Erstellung der Präferenzordnung verbundenen Rechenaufwand erheblich reduzieren. Von besonderer Bedeutung ist dabei die Transformation in eine Opportunitätskostentabelle, die mindestens ebensoviele Nullen wie Spalten enthält und in der zudem alle Werte positiv sind.

Die Entscheidungsbaum-Darstellung eignet sich besonders für Spezialfälle des allgemeinen Entscheidungsmodells[3]. Ein Entscheidungsbaum besteht aus Aesten und zwei Arten

---

[1] Jeffrey, S. 27

[2] Zur Definition des Erwartungswertes vgl. S. 53 ff. Den späteren Ausführungen etwas vorgreifend, wird hier bereits eine lineare Nutzenfunktion vorausgesetzt. Vgl. dazu S. 11

[3] Vgl. etwa Nievergelt, S. 101 ff.

von Knoten. Ein quadratischer Knoten heisst "Entscheidungsknoten" und bezeichnet die Möglichkeit eines Aktors, einen Entscheid zu fällen. Die diesem Knoten folgenden Aeste symbolisieren die Handlungsmöglichkeiten oder Aktionen des Aktors. Ein Kreis-Knoten wird "Zufallsknoten" genannt; die Aeste, die von ihm ausgehen, zeigen die möglichen Zustände der Umwelt. Das Ende eines jeden Astes wird mit einem Punkt markiert, dem die Konsequenzen oder Opportunitätskosten der entsprechenden Zugfolge zugeschrieben werden (vgl. Abb. 1).

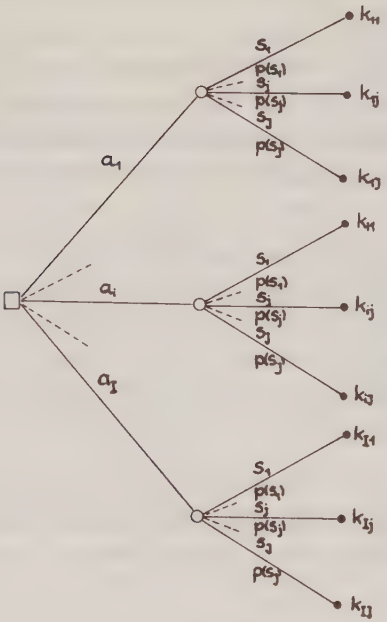


Abb. 1: Darstellung des allgemeinen Entscheidungsmodelles mit einem Entscheidungsbaum

## 3. Klassifikation von Entscheidungsmodellen[1]

Grundsätzlich können sämtliche konstitutiven Elemente des allgemeinen Entscheidungsmodells und Kombinationen dieser Elemente als Klassifikationsmerkmale für die Klassifikation von Entscheidungsmodellen verwendet werden. Im folgenden werden die wichtigsten nach diesen Gesichtspunkten gebildeten Klassen aufgezählt:

- Bzgl. des Aktors: Modelle der Individualentscheidung; Modelle der Gruppenentscheidung.
- Bzgl. der Anzahl Zielsetzungen: Modelle mit einem Ziel; Modelle mit mehreren Zielen.
- Bzgl. der Anzahl Aktionen: Modelle mit zwei Aktionen (2-Aktionen-Probleme); Modelle mit mehreren Aktionen.
- Bzgl. der Anzahl Umweltzustände: Modelle mit nur einem Umweltzustand (Sicherheitssituation); Modelle mit mehreren Umweltzuständen (Unsicherheitssituation).
- Bzgl. der Wahrscheinlichkeiten: Modelle mit bekannter Wahrscheinlichkeitsverteilung über dem (natürlich mehr als ein Element umfassenden) Zustandsraum (Risikosituation oder Situation der stochastischen Unsicherheit[2]); Modelle ohne Wahrscheinlichkeitsverteilung über dem Zustandsraum (Ungewissheitssituation oder Situation der rationalen Indeterminiertheit[3]).
- Bzgl. der Konsequenzen: Unterscheidung von Modellen nach dem Niveau der Konsequenzenmessung.
- Bzgl. der Entscheidungsmaxime: Modelle mit nicht-probabilistischen Entscheidungsmaximen; Modelle mit probabilisti-

---

[1] Vgl. etwa Bamberg/Coenenberg, S. 35 f.

[2] Kühn, S. 32

[3] Kühn, S. 35

schen Entscheidungsmaximen. Die probabilistische Entscheidungsmaxime ist die bereits erwähnte Erwartungswertmaxime. Entscheidungsmodelle, die auf der Erwartungswertmaxime aufbauen, heissen auch "Bayes'sche Modelle"[1].

Gelegentlich wird die Unsicherheitssituation noch nach dem Kriterium der Häufigkeit der Wiederholung einer Entscheidung unterteilt. Quasisicherheit oder stochastische Sicherheit liegt dann vor, wenn eine Entscheidung in der Vergangenheit schon sehr oft getroffen wurde, so dass die Wahrscheinlichkeitsverteilung durch relative Häufigkeiten bestimmt wird[2] und der gleiche Entscheid auch in Zukunft sehr oft gefällt werden wird[3]. Von "stochastischer Unsicherheit" spricht man nach dieser Interpretation nur, wenn diese Bedingungen nicht erfüllt sind.

Dem so umschriebenen Fall der Quasisicherheit dürfte allerdings in der Realität kaum Bedeutung zukommen[4]. Ein Indiz für die Geltung dieser Behauptung ist das Fehlen dieses Entscheidungstyps bei sehr vielen Gliederungen der Literatur, insbesondere in derjenigen aus dem angelsächsischen Sprachbereich[5].

## 4. Einführung in das Bayes'sche Modell des Entscheidens

Das Bayes'sche Modell kann in der Klassifikation, die von den konstituierenden Merkmalen des allgemeinen Entschei-

[1] Jeffrey, S. 9

[2] Vgl. dazu den Exkurs: Deutungen des Wahrscheinlichkeitsbegriffs, S. 29 f.

[3] Kühn, S. 31

[4] Vgl. Kühn, S. 31 f.

[5] Z.B. bei Ewart/Fort/Lin, S. 250 ff.; Luce/Raiffa, S. 13; Miller/Starr, S. 80 f.

dungsmodells ausgeht, der Unsicherheitssituation zugeordnet werden, d.h. es enthält immer mehrere Umweltzustände. Zudem geht ein Verfechter des Bayesianischen Modells (ein sogenannter "Bayesianer") davon aus, dass immer eine Wahrscheinlichkeitsverteilung über dem Zustandsraum bekannt sei, die Unterscheidung des Unsicherheitsfalles in eine Risiko- und eine Ungewissheitssituation erscheint ihm künstlich; er akzeptiert nur die Risikosituation. Den Fall der rationalen Indeterminiertheit lehnt er als nicht realistisch ab[1].

Schliesslich werden im Bayes'schen Modell die Konstituanten des allgemeinen Entscheidungsmodells durch ein weiteres Merkmal ergänzt: durch die Möglichkeit nämlich, zusätzliche Information über die Umweltzustände einzuholen und dadurch die ursprüngliche Wahrscheinlichkeitsverteilung über dem Zustandsraum - die Apriori-Verteilung - im Lichte der erweiterten empirischen Kenntnis nach dem Prinzip des Lernens aus der Erfahrung zu revidieren[2]. Die so erhaltenen Wahrscheinlichkeiten heissen "Aposteriori-Wahrscheinlichkeiten". Dabei unterscheiden sich die beiden Termini "Apriori" und "Aposteriori" nur relativ in bezug auf den Zeitpunkt der Informationserhebung; grundsätzlich können Aposteriori-Wahrscheinlichkeiten beliebig oft wieder als Apriori-Werte verwendet werden[3].

In Abb. 2 wird das Bayes'sche Modell mit einem Entscheidungsbaum dargestellt. Dabei symbolisieren

---

[1] Winkler, S. 221

[2] Vgl. Stegmüller (4), S. 124

[3] Schlaifer, S. 196 f.

$E = \{e_1,\ldots,e_n,\ldots,e_N\}$: Die Menge aller zur Verfügung stehender Möglichkeiten $e_n$, $n=1,2,\ldots,N$, Informationen zu erheben (Experimente);

$R_n = \{r_1^n,\ldots,r_k^n,\ldots,r_k^n\}$: Die Menge der möglichen Resultate $r_k^n$, $k=1,2,\ldots,K$ eines Experimentes $e_n$;

$R = \bigcup_n R_n$: Die Menge aller möglicher Resultate sämtlicher zur Verfügung stehender Experimente.

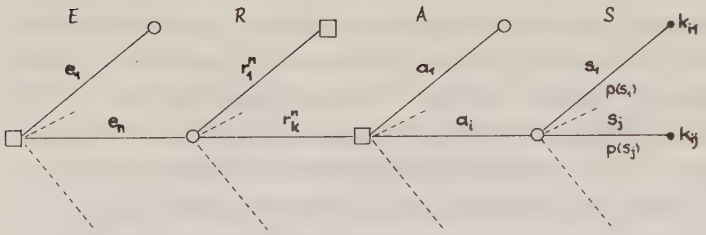


Abb. 2: Das Bayes'sche Modell des Entscheidens

In der Bayes'schen Entscheidungssituation kann der Aktor nicht - wie beim allgemeinen Modell - bloss eine optimale Aktion wählen: er hat sich zunächst für eines der zur Verfügung stehenden Experimente zu entscheiden und braucht zudem noch eine Regel, welche jedem möglichen Resultat des Experimentes die optimale Aktion zuordnet. Eine solche Regel heisst "Entscheidungsregel"; Die Bezeichnung eines Experimentes und der dazugehörenden Entscheidungsregel wird "Strategie" genannt[1]. Eine Strategie enthält also für jeden Entscheidungsknoten eine Anweisung über die zu treffende Entscheidung. Der hier betrachtete Aktor hat zwischen den verschiedenen möglichen Strategien seine Wahl zu treffen.

[1] Raiffa/Schlaifer, S. 11 f.

Bei der Besprechung der "Klassifikation von Entscheidungsmodellen" wurde bereits auf die Verbindung der Erwartungswertmaxime mit dem Namen "Bayes" hingewiesen[1]. Der Bayesianer trifft in der Tat alle seine Entscheidungen aufgrund der Erwartungswertmaxime, die - nach dem englischen Mathematiker Thomas Bayes - auch "Bayes'sches Prinzip" genannt wird[2]. Ganz allgemein besagt dieses Prinzip bloss, dass irgendwelche Grössen mit den korrespondierenden Wahrscheinlichkeiten gewichtet werden[3]. Ueber die Natur dieser Grössen hat sich bekanntlich ein sehr wesentlicher Teil der Diskussion im Rahmen der Risikoproblematik abgespielt und ihren Niederschlag im Nutzenkonzept gefunden[4]. Dieser Aspekt wird in dieser Arbeit nicht berührt; es wird davon ausgegangen, dass Risikoneutralität (oder eine lineare Nutzenfunktion[5]) vorliege, was eine Transformation von Auszahlungen in Nutzenwerte erübrigt.

Um den Erwartungswert einer Strategie berechnen zu können, muss der Aktor all jenen Aesten des Entscheidungsbaumes, die von den Zufallsknoten ausgehen, Wahrscheinlichkeiten zuordnen können. Wie aus Abb. 2 ersichtlich, handelt es sich dabei um die bedingten Wahrscheinlichkeiten $p(r_k^n/e_n)$ und $p''(s_j/r_k^n,e_n)$. Der zweite Ausdruck bezeichnet die Aposteriori-Wahrscheinlichkeit, die hier - wie vielerorts üblich - mit zwei Strichen ausgezeichnet wird (während die Apriori-Wahrscheinlichkeit mit einem Strich notiert wird).

In der Literatur werden drei Basis-Modelle unterschieden,

---

[1] Vgl. S. 10

[2] Jeffrey, S. 9

[3] Vgl. Bamberg/Coenenberg, Fussnote S. 78

[4] Vgl. Bamberg/Coenenberg, S. 62 ff.

[5] Vgl. Bamberg/Coenenberg, S. 71 ff.

mit deren Hilfe die gesuchten Wahrscheinlichkeiten ermittelt werden können[1] (zur Illustration sei ein einfaches Zahlenbeispiel verwendet):

1. Es sind die gemeinsamen Wahrscheinlichkeiten $p(s_j, r_k^n / e_n)$ definiert auf der Menge $S \times R$ gegeben:

| | $s_1$ | $s_2$ |
|---|---|---|
| $r_1^n$ | 0.54 | 0.06 |
| $r_2^n$ | 0.225 | 0.075 |
| $r_3^n$ | 0.03 | 0.07 |

Die bedingten Wahrscheinlichkeiten $p(r_k^n / e_n)$ entsprechen in diesem Falle den marginalen Wahrscheinlichkeiten auf $R$.

$$(1) \quad p(r_k^n / e_n) = \sum_{j=1}^{J} p(s_j, r_k^n / e_n), \qquad k = 1, \ldots, K$$

d.h. $p(r_1^n / e_n) = 0.6;$

$p(r_2^n / e_n) = 0.3;$

$p(r_3^n / e_n) = 0.1.$

Die Aposteriori-Wahrscheinlichkeiten errechnen sich aufgrund der Beziehung

$$(2) \quad p''(s_j / r_k^n, e_n) = \frac{p(s_j, r_k^n / e_n)}{p(r_k^n / e_n)}$$

und man erhält auf diese Weise folgende Tabelle von

[1] Raiffa/Schlaifer, S. 5

Aposteriori-Wahrscheinlichkeiten:

| | $s_1$ | $s_2$ |
|---|---|---|
| $r_1^n$ | 0.9 | 0.1 |
| $r_2^n$ | 0.75 | 0.25 |
| $r_3^n$ | 0.3 | 0.7 |

2. Es sind die marginalen Wahrscheinlichkeiten auf $R$, in der entscheidungstheoretischen Terminologie die Apriori-Wahrscheinlichkeiten $p's(s_j)$,

$$p'(s_1) = 0.795,$$

$$p'(s_2) = 0.205$$

und die bedingten Wahrscheinlichkeiten $p(r_k^n/s_j,e_n)$ gegeben:

| | $s_1$ | $s_2$ |
|---|---|---|
| $r_1^n$ | 0.679 | 0.293 |
| $r_2^n$ | 0.283 | 0.366 |
| $r_3^n$ | 0.038 | 0.341 |

Aufgrund der Beziehung

$$(3) \quad p(s_j,r_k^n/e_n) = p'(s_j) \cdot p(r_k^n/s_j,e_n)$$

kann die Tabelle mit den gemeinsamen Wahrscheinlichkeiten $S \times R$ aufgestellt werden. Die gesuchten Wahrscheinlichkeiten können nun wie bei Fall 1 ermittelt werden.

3. Es sind die bedingten Wahrscheinlichkeiten $p(r_k^n/e_n)$ bzw. die marginalen Wahrscheinlichkeiten auf $R$

$$p(r_1^n/e_n) = 0.6$$

$$p(r_2^n/e_n) = 0.3$$

$$p(r_3^n/e_n) = 0.1$$

und die Aposteriori-Wahrscheinlichkeiten $p''(s_j/r_k^n,e_n)$ gegeben. Die Tabelle mit den gemeinsamen Wahrscheinlichkeiten $S \times R$ kann aufgrund der Beziehung (2) aufgestellt werden. Die marginalen Wahrscheinlichkeiten auf $S$, $p'(s_j)$, ergeben sich aufgrund der Beziehung (4) aus dieser Tabelle:

$$(4) \quad p'(s_j) = \sum_{j=1}^{K} p(s_j,r_k^n/e_n), \qquad j=1,\ldots,J.$$

Zusammenfassend lassen sich die drei Methoden folgendermassen charakterisieren:

Mit Hilfe der Beziehungen (1) - (4) können sämtliche Elemente der Menge $\{p'(s_j), p(r_k^n/e_n), p(s_j,r_k^n/e_n), p(r_k^n/s_j,e_n), p''(s_j/r_k^n,e_n)\}$ hergeleitet werden, wenn entweder

1. die $p(s_j,r_k^n/e_n)$ oder
2. die $p'(s_j)$ und die $p(r_k^n/s_j,e_n)$ oder
3. die $p(r_k^n/e_n)$ und die $p''(s_j/r_k^n,e_n)$

gegeben sind.

Welche dieser drei Methoden in einem konkreten Fall angewendet wird, hängt vom Wissen des Aktors, von seiner Erfahrung ab: er wird diejenigen Wahrscheinlichkeiten als gegeben annehmen, die ihm erlauben, den besten Gebrauch von dieser Erfahrung zu machen[1]. Bestimmt wird dieses Wissen vor allem auch durch die Art der zusätzlichen Information. In der wohl berühmtesten und meistzitierten Fallstudie zur Bayes'schen Entscheidungstheorie, dem Erdöl-Beispiel[2], wird zusätzliche Information mittels seismologischer Experimente, die Aufschluss über die geologische Struktur eines Gebietes geben sollen, erhoben. Das Wissen um diese Struktur lässt auf die Wahrscheinlichkeit von Erdölvorkommen im untersuchten Gebiet schliessen. In dieser Fallstudie wird davon ausgegangen, dass der Aktor aufgrund seiner Erfahrung die Aposteriori-Wahrscheinlichkeiten $p"(s_j/r_k^n, e_n)$ und die Wahrscheinlichkeit, dass im Gebiet die geologische Struktur x vorkommt - $p(r_k^n)$ -, direkt bestimmen kann. Die für die Berechnung der Erwartungswerte der möglichen Strategien benötigten Wahrscheinlichkeiten können hier offensichtlich mit Hilfe der Methode 3 ermittelt werden.

Aufgrund der bisherigen Ausführungen scheint Methode 3 überflüssig zu sein, denn die hier als für die Lösung notwendig deklarierten Wahrscheinlichkeiten sind ja schon bekannt. Nun wurde bisher aber eine mögliche Strategie unterschlagen: diejenige nämlich, gar kein Experiment durchzuführen und direkt die optimale Aktion zu bestimmen. Dieses "Nicht-Experiment" wird in der Regel mit $e_o$ bezeichnet und der Ast des Entscheidungsbaumes, der mit $e_o$ beginnt, hat folgendes Aussehen:

[1] Raiffa/Schlaifer, S. 5

[2] Raiffa/Schlaifer, S. 20 ff.; Feichtinger, S. 453 ff.

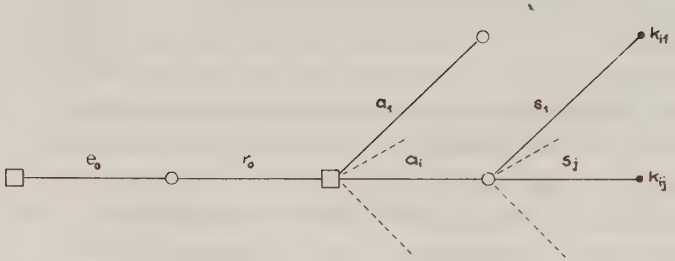


Abb. 3: Entscheidungsbaum der $e_o$-Strategie

Aus Abb. 3 wird ersichtlich, dass der Aktor für die Berechnung des Erwartungswertes der $e_o$-Strategie die Wahrscheinlichkeiten $p'(s_j)$ - die Apriori-Wahrscheinlichkeiten - benötigt, die er eben mit Hilfe der Methode 3 ermitteln kann.

Eine andere wichtige Technik, zusätzliche Informationen zu gewinnen, besteht in der Durchführung von Stichproben. Diese Arbeit befasst sich ausschliesslich mit dieser Problematik, und es wird im nächsten Teil dargelegt werden, dass und warum der Aktor in diesem Falle über Kenntnis der bedingten Wahrscheinlichkeiten $p(r_k^n/s_j,e_n)$ verfügt. Zudem wird in der Regel davon ausgegangen, dass die Apriori-Verteilung über dem Zustandsraum gegeben sei. Die für den Entscheid benötigten Wahrscheinlichkeiten können also mit Methode 2 ermittelt werden. Hier soll diese Situation, in der der Aktor die $p'(s_j)$ und $p(r_k^n/s_j,e_n)$ kennt, "Normalkonstellation der Wahrscheinlichkeiten" genannt werden und für den weiteren Verlauf dieser Arbeit wird davon ausgegangen, dass immer diese Konstellation vorliege.

An dieser Stelle können zwei für die Bayes'sche Entscheidungstheorie wesentliche Termini eingeführt werden: die bedingte Wahrscheinlichkeit $p(r_k^n/s_j,e_n)$ heisst "Likelihood" bzw. $f(r_k^n/s_j,e_n)$ "Likelihoodfunktion"[1]. Die mit den Apriori-Wahrscheinlichkeiten gewichteten Likelihoods ermöglichen nach Beziehung (3) die verbundenen Wahrscheinlichkeiten auf $S \times R$ zu ermitteln und die Relation (2) kann folgendermassen ausgedrückt werden:

$$(5) \quad p''(s_j/r_k^n,e_n) = \frac{p'(s_j) \cdot p(r_k^n/s_j,e_n)}{p(r_k^n/e_n)}$$

Die Beziehung (5) heisst "Bayes-Theorem". Weil

$$p(r_k^n/e_n) = \sum_{j=1}^{J} p(r_k^n/s_j,e_n) \cdot p'(s_j), \qquad (1.1)$$

wird das Bayes-Theorem oft folgendermassen geschrieben:

$$p''(s_j/r_k^n,e_n) = \frac{p'(s_j) \cdot p(r_k^n/s_j,e_n)}{\sum_{j=1}^{J} p(r_k^n/s_j,e_n) \cdot p'(s_j)} \qquad (1.2)$$

Zum zweiten Mal gelangt man hier mit dem Namen von Thomas Bayes in Berührung, und wie die Erwartungswertmaxime ist auch das Bayes-Theorem - an sich nichts anderes als eine "triviale Verknüpfung bedingter Wahrscheinlichkeiten"[2] - ein zentrales Element, der mit dem Namen des berühmten Mathematikers gekennzeichneten Entscheidungstheorie.

Der Bayesianer bezieht auch die Kosten der Informationsbeschaffung in seinen Kalkül ein: Grundsätzlich kann er das tun, indem er entweder vom Konsequenzenerwartungswert jeder Strategie die Kosten des Experimentes abzieht oder

[1] Bamberg/Coenenberg, S. 114

[2] Schaefer, S. 130

die im Entscheidungsbaum erscheinenden Konsequenzenwerte um die entsprechenden Kosten korrigiert.

Die hier in ihren Grundzügen geschilderte Entscheidung heisst "Praeposteriori-Entscheidung", weil ein Entscheid gefüllt wird, bevor ein Experiment überhaupt durchgeführt wird. Eine derartige Entscheidung basiert auf den potentiellen Resultaten von Experimenten und potentiellen Aposteriori-Verteilungen[1]. Im Gegensatz dazu wird bei der "Aposteriori-Entscheidung" davon ausgegangen, dass ein Experiment bereits durchgeführt wurde und das Resultat bekannt ist. Der Aktor, der eine Praeposteriori-Entscheidung fällt, wählt also eine Strategie, derjenige, der eine Aposteriori-Entscheidung zu treffen hat, direkt eine Aktion.

## 5. Prämissen

Zwei für den weiteren Verlauf der Arbeit entscheidende Prämissen wurden bereits formuliert[2]:

- Das Experiment ist eine Zufallsstichprobe: Aus einer statistischen Menge (Grundgesamtheit) wird eine endliche Menge so entnommen, dass jedes Element der Grundgesamtheit die gleiche Chance hat, ausgewählt zu werden[3]. Unter "zusätzlicher Information" wird deshalb immer und ausschliesslich Stichprobeninformation verstanden.
- Es wird von der Voraussetzung der Risikoneutralität ausgegangen. Eine Transformation von Geldwerten (Konsequenzen oder Opportunitätskosten) in Nutzen wird nicht durchgeführt.

[1] Winkler, S. 297

[2] Vgl. S. 13 und S. 18

[3] Hengst, S. 51

Ferner werden im folgenden ausschliesslich 2-Aktionen-Probleme, die einen besonders universellen Entscheidungstyp - die Entscheidung "etwas tun oder nicht-tun" - repräsentieren, behandelt. Einige der Methoden, die hier für die Lösung von Entscheidungsproblemen entwickelt werden, können ausschliesslich auf 2-Aktionen-Probleme angewandt werden. Es wird aber immer ein allgemeiner Ansatz geboten, der auch die Lösung von Mehr-Aktionen-Problemen ermöglicht.

## TEIL II EINE TYPOLOGIE VON BAYES-ENTSCHEIDUNGEN[1]

Eine Typologie ist eine sinnvolle Kombination von Merkmalen eines Untersuchungsgegenstandes, so dass ein "Typus, das für einen bestimmten Untersuchungszweck Wesentliche, Charakteristische einer Anzahl von Objekten wider-(spiegelt)"[2]. Die Kriterien, denen eine Typologie genügen soll, hangen vom Untersuchungszweck ab. Hier erschienen die folgenden beiden Kriterien von Bedeutung: Erstens ein didaktisches: die Typenbildung soll ein Fortschreiten vom Einfacheren zum Schwierigeren ermöglichen; zweitens ein methodisches: die Typen sollen sich in bezug auf die Methoden, die für die Lösung der Probleme innerhalb eines Types verwendet werden, unterscheiden.

Die folgenden typologischen Merkmale erscheinen für eine diesen Kriterien genügende Typenbildung geeignet:

- Die Informationsstruktur;
- das stochastische Modell;
- die Apriori-Verteilung;
- die Art der Konsequenzenfunktion.

Das didaktische Kriterium hat nun allerdings zur Folge, dass einige der hier zu entwickelnden und zu behandelnden Typen gar nicht als eigentliche Bayes-Entscheidungen bezeichnet werden können. Das gilt insbesondere für die Entscheidung ohne zusätzliche Information, denn zusätzliche Information ist ja konstitutives Element einer Bayes-Entscheidung[3]. Beschränkt gilt das auch für Entscheidungen

[1] Der Begriff "Bayes-Entscheidung" wurde im vorgehenden Teil nicht verwendet; es sei darunter eine auf dem Bayes'schen Modell aufbauende Entscheidung verstanden.

[2] Knoblich, S. 143

[3] Vgl. S. 11

mit zusätzlicher Information, aber gegebenem Informationsumfang. In diesem Falle hat ein Aktor bloss unter Entscheidungsregeln zu wählen, während ein Bayesianer in der Regel Entscheide über Strategien fällt.

Zweck dieser "didaktischer Kategorien" ist es, ein sukzessives Vorarbeiten zu den eigentlichen Bayes-Entscheidungen zu ermöglichen.

## 1. Typologische Merkmale

### 1.1. Die Informationsstruktur

Beim Merkmal "Informationsstruktur" werden die folgenden Merkmalsausprägungen berücksichtigt:

- Keine zusätzliche Informationsaktivität;
- zusätzliche Informationsaktivität;
  - gegebener Informationsumfang;
  - variabler Informationsumfang;
    - einmaliger Informationsentscheid;
    - mehrstufiger (sequentieller) Informationsentscheid.

### 1.2. Das stochastische Modell

Die Informationsaktivität besteht in der Beobachtung von Zufallsvariablen, die von einem Zufallsexperiment - der Stichprobe - erzeugt werden. Wenn der Aktor diesen Prozess (antizipativ) in seinem Entscheidungskalkül mitberücksichtigen will, braucht er eine Vorstellung - präziser: ein Modell - dieses Prozesses, damit er die, wie im Abschnitt "Grundzüge der Bayes'schen Entscheidungstheorie" gezeigt, zentrale Likelihoodfunktion ermitteln kann. Ein solches Modell heisst "stochastisches Modell" und beinhaltet einen Komplex von Voraussetzungen sowie eine "theoretische Verteilung", die angibt, "wie oft ein bestimmtes Ereignis

bzw. die ihm zugeordnete Zufallsvariable bei fortlaufender Realisation des Modells auftritt"[1/2].

Das stochastische Modell dient also dem Aktor dazu, einen Zufallsprozess, der in Zukunft stattfinden wird, abzubilden - der Begriff "Modell" deutet die damit verbundene Abstraktion an. Ein Modell kann nicht richtig oder falsch sein, nur "mehr oder weniger angemessen"[3]; Menges spricht davon, dass die mit einem Verteilungstyp verbundenen Voraussetzungen vom Prozess, der abgebildet werden soll, "wenigstens approximativ" erfüllt werden müssen[4]. Letztlich ist aber die Annahme, dass sich dieser Prozess verhalte wie ein bestimmtes stochastisches Modell, subjektiv[5].

"Die meisten stochastischen Modelle lassen sich ... von alters her durch ein Urnenschema abbilden, bei dem aus einem System passend kombinierter Urnen, die nach bestimmten Vorschriften mit verschiedenfarbigen Kugeln gefüllt sind, sukzessive eine bestimmte Zahl von Kugeln nach vorgegebenen Regeln gezogen werden"[6].

Im folgenden sollen die in der entscheidungstheoretischen Literatur am häufigsten verwendeten Modelle kurz dargestellt werden.

### 1.21. Das Binomial-Modell

Voraussetzungen:

a) Die Anzahl n der Versuche ist fest gegeben.

b) Jeder Versuch hat nur zwei mögliche Ergebnisse. (Einem

---

[1] Hengst, S. 58

[2] In der angelsächsischen Literatur wird in diesem Zusammenhang der Ausdruck "data generating process" verwendet. S. bei Raiffa/Schlaifer, S. 36; Winkler, S. 149

[3] Kriz, S. 25

[4] Menges (1), S. 226

[5] Winkler, S. 82

[6] Hengst, S. 59

weit verbreiteten Gebrauch folgend, soll hier ebenfalls vom "günstigen" bzw. "ungünstigen" Ergebnis die Rede sein.)

c) Die Wahrscheinlichkeit $\theta$ des Auftretens des günstigen Ergebnisses bzw. $(1-\theta)$ des ungünstigen Ergebnisses ist bei jedem Versuch konstant.

d) Die Versuche sind voneinander unabhängig.

Eine binomial-verteilte Zufallsvariable R hat die Massefunktion

$$p(R=r;n,\theta) = \begin{cases} \binom{n}{r}\theta^r (1-\theta)^{n-r} & \text{für } r=0,1,\ldots,n \\ 0 & \text{sonst,} \end{cases} \quad (2.1)$$

mit n und $\theta$ als Parameter[1].

Dem Binomial-Modell entspricht das Bild einer mit zwei Arten von Kugeln gefüllten Urne, aus der n-mal eine Kugel gezogen und die gezogene Kugel immer wieder zurückgelegt wird.

## 1.22. Das hypergeometrische Modell

Dieses Modell unterscheidet sich vom Binomial-Modell durch die Aufhebung der Voraussetzungen c) und d). Auf das Urnenschema bezogen, heisst das, dass die jeweils gezogene Kugel nicht mehr zurückgelegt wird.

Eine hypergeometrisch verteilte Zufallsvariable R hat die

[1] Eine Zufallsvariable wird jeweils mit grossen, ihre Realisation mit kleinen lateinischen Buchstaben bezeichnet.

Massefunktion

$$p(R=r;n,\theta,N) = \begin{cases} \dfrac{\binom{\theta N}{r}\binom{(1-\theta)N}{n-r}}{\binom{N}{n}} & \text{für } r=0,1,\ldots,n \\ 0 & \text{sonst.} \end{cases} \qquad (2.2)$$

Zu den beiden Parametern des Binomial-Modells kommt als dritter Parameter N, die Grösse der Grundgesamtheit (im Urnenmodell: Gesamtheit der Kugeln) dazu.

1.23. Das Poisson-Modell

Das Poisson-Modell kann auf zwei verschiedene Arten hergeleitet bzw. interpretiert werden. Entweder als Grenzwert des Binomial-Modells, wenn

$$\theta \longrightarrow 0,\ n \longrightarrow \infty \text{ und } n\theta \longrightarrow \lambda.$$

Eine poisson-verteilte Zufallsvariable hat dann die Massefunktion

$$p(R=r;\lambda) = \begin{cases} \dfrac{(\lambda)^r}{r!} e^{-\lambda} & \text{für } r=0,1,\ldots,\infty \\ 0 & \text{sonst.} \end{cases} \qquad (2.3)$$

(Faustregel: Wenn $n\theta=\lambda\leq 5$ und $n>20$, kann das Poisson-Modell als Annäherung des Binomial-Modells verwendet werden[1].

[1] Yamane, S. 738

Das Poisson-Modell kann aber auch über den Poisson-Prozess hergeleitet werden. Die Charakteristik dieses Prozesses ist das Auftreten der Ereignisse über einem Zeitkontinuum. Dabei wird vorausgesetzt, dass die Ereignisse zufällig auftreten (d.h., dass die Wahrscheinlichkeit des Auftretens eines Ereignisses in einem Zeitintervall konstant bleibt) und die Ereignisse eines Zeitintervalls unabhängig von denjenigen eines anderen Intervalls sind.

Im Unterschied zum Binomial-Modell ist beim Poisson-Modell die Anzahl n der Versuche unbekannt. Wenn beim Poisson-Modell von r günstigen Ereignissen bzw. Erfolgen die Rede ist, so meint man r Erfolge in einem gegebenen Intervall t[1].

Die Wahrscheinlichkeit des Auftretens von r Erfolgen im Intervall t ist

$$p(R=r;\lambda t) = \begin{cases} \frac{(\lambda t)^r}{r!} e^{-\lambda t} & \text{für } r=0,1,\ldots \\ 0 & \text{sonst,} \end{cases} \qquad (2.4)$$

wobei $\lambda$ die Anzahl Erfolge in einer Einheit des Intervalls symbolisiert. $\lambda t$ ist der einzige Parameter dieses Modells.

## 1.24. Das Normal-Modell

Die Zufallsvariablen der bisher vorgestellten Modelle waren alle diskret. Die normalverteilte Zufallsvariable hingegen ist eine stetige Zufallsvariable X, die ursprünglich mit Hilfe des Grenzübergangs $n \longrightarrow \infty$ vom Binomial-

[1] Yamane, S. 744

Modell gewonnen wurde (De-Moivre-Laplacescher Grenzübergang)[1]. Die Dichtefunktion der normalverteilten Zufallsvariablen X ist

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.5}$$

Die beiden Parameter sind $\mu$, der Mittelwert und $\sigma$, die Standardabweichung der Zufallsvariablen X. Eine (0,1)-normalverteilte Zufallsvariable heisst "standardnormalverteilt" und hat die Dichtefunktion

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \tag{2.6}$$

wobei

$$Z = \frac{X-\mu}{\sigma} \tag{2.7}$$

Die Tatsache, dass das Normal-Modell historisch gesehen vom Binomial-Modell abstammt, deutet schon auf eine Verwendungsmöglichkeit des Modells, die Approximation anderer Modelle, hin. Von Bedeutung ist dies vor allem deshalb, weil die bisher behandelten diskreten Modelle bei grossen Stichprobenumfängen rechnerisch nur noch schwer zu bewältigen sind.

Das Binomial-Modell lässt sich durch das Normal-Modell mit dem Erwartungswert $\theta$ und der Standardabweichung

$$\sqrt{\frac{\theta(1-\theta)}{n}}$$

approximieren, wenn $n\cdot\theta\cdot(1-\theta) \geq 9$[2].

[1] Vgl. Menges (1), S. 238 ff.

[2] Sachs, S. 166

Mit dem gleichen Normal-Modell kann auch das hypergeometrische Modell approximiert werden, wenn die Bedingung $n\Theta \geq 4$ erfüllt ist[1].

Wenn $\lambda t \geq 9$, kann schliesslich auch das Poisson-Modell durch ein Normal-Modell mit dem Erwartungswert $\lambda t$ und der Standardabweichung $\sqrt{\lambda t}$ angenähert werden[2].

Die zentrale Bedeutung des Normal-Modells rührt aber vom "zentralen Grenzwertsatz" her:

"Das Modell, das diesen Satz so eminent wichtig für die Praxis werden lässt, ist das folgende:

1. Gegeben ist eine Erscheinung, die von einem Komplex zahlreicher Faktoren beeinflusst, beherrscht, determiniert wird.
2. Jeder einzelne Faktor des Komplexes übt nur eine sehr kleine Wirkung aus; für sich allein genommen ist seine Rolle praktisch vernachlässigbar.
3. Jedem einzelnen Faktor wird eine Zufallsvariable $X_i$ $(i=1,2,\ldots,n)$ zugeordnet.
4. Es ist nicht bekannt, wieviel Zufallsvariablen (und damit Faktoren) beteiligt sind.
5. Es ist nicht bekannt, wie die einzelnen $X_i$ verteilt sind.
6. Man weiss indessen (oder nimmt an), dass die $X_i$ (und damit die Faktoren) sich gegenseitig nicht beeinflussen.

... Unter den getroffenen Annahmen ist die Summe $X=X_1+\ldots X_n$ (der Totaleffekt aller $X_i$ auf das betreffende Phänomen, d.V.) normalverteilt"[3].

Die Anwendung des zentralen Grenzwertsatzes in der Stichprobentheorie führt zu folgenden Ergebnissen[4]:

- Bei einer $(\mu,\sigma)$normalverteilten, grossen Grundgesamtheit sind die stichprobenmittelwerte $(\mu, \frac{\sigma}{\sqrt{n}})$ normal-

[1] Sachs, S. 175

[2] Sachs, S. 187

[3] Menges (1), S. 249 f.

[4] Yamane, S. 162 ff.

verteilt.

- Auch wenn die (ebenfalls grosse) Grundgesamtheit nicht normalverteilt ist, sin die Stichprobenmittelwerte näherungsweise $(\mu, \frac{\sigma}{\sqrt{n}})$ normalverteilt, allerdings nur, wenn $n \geq 30$.

## 1.3. Die Apriori-Verteilung

Die Apriori-Verteilung nimmt - als konstitutives Element des Bayes'schen Modells des Entscheidens - eine zentrale Stellung in der Bayes'schen Entscheidungstheorie ein. Dieser Abschnitt befasst sich zunächst mit Interpretationen der Apriori-Verteilung, dann mit der Rolle der Apriori-Verteilung in der Bayes'schen Entscheidungstheorie und schliesslich mit den verschiedenen Ausprägungen des typologischen Merkmals "Apriori-Verteilung".

Eine Apriori-Verteilung interpretieren, heisst eine Aussage über den Charakter der Apriori-Wahrscheinlichkeiten treffen. Zu diesem Zweck wird dem Abschnitt über die Interpretation der Apriori-Verteilung ein vorbereitender Exkurs über Deutungen des Wahrscheinlichkeitsbegriffs vorangestellt.

Exkurs: Deutungen des Wahrscheinlichkeitsbegriffs

Wer sich mit dem Begriff der Wahrscheinlichkeit auseinandersetzt, wird mit einer Fülle von verschiedenartigen, sich teils überschneidenden, teils konkurrierenden Definitionen und Interpretationen konfrontiert. Es wird hier versucht, einen Ueberblick über die verschiedenen Positionen zu geben. Anschliessend soll der in dieser Arbeit vertretene Standpunkt umrissen werden.

Heute ist wohl unbestritten, dass die axiomatische Charakterisierung des Wahrscheinlichkeitsbegriffs die Grundlage jeder Theorie der Wahrscheinlichkeit darstellt:

"Der Begriff der Wahrscheinlichkeit wird als undefinierter Grundbegriff eingeführt ... Dies bedeutet: Die möglichen Deutungen von 'Wahrscheinlichkeit' werden nur einer formalen Einschränkung unterworfen"[1].

Die anerkannte Axiomatik wurde vom russischen Wahrscheinlichkeitstheoretiker A.N. Kolmogoroff formuliert, nach dem die Axiome auch benannt sind[2].

Stegmüller hat darauf hingewiesen, dass durch diese Axiomatik wohl die formalen Eigenschaften der Wahrscheinlichkeit festgelegt werden, dass aber die folgenden Fragen offen bleiben:

"(1) Welche 'vollständige' Interpretation des Wahrscheinlichkeitskalküls ist als inhaltlich adäquat anzusehen?
(2) Wie gelangt man ... zu Wahrscheinlichkeitsaussagen?
(3) Wie überprüft man Wahrscheinlichkeitsbehauptungen?
(4) Wie lauten die Regeln für die korrekte Anwendung des Wahrscheinlichkeitsbegriffs ...?"[3]

Vor allem in bezug auf die Beantwortung der Fragen (1)

[1] Stegmüller (3), S. 107

[2] Vgl. Menges (1), S. 29 f. und 85 ff.

[3] Stegmüller (3), S. 109

und (2) können drei verschiedene Wahrscheinlichkeitsbegriffe - logischer, objektiver und subjektiver - und damit verbundene Konzeptionen unterschieden werden, die in der Literatur eine lange Tradition haben. Wir wollen sie "klassische Wahrscheinlichkeitskonzepte" nennen. Bevor auf sie eingegangen wird, soll eine andere Kategorie von Definitionen, die sogenannten Nominaldefinitionen des Terminus Wahrscheinlichkeit, kurz vorgestellt werden:

a) Nominaldefinitionen der Wahrscheinlichkeit

"Unter einer Nominaldefinition versteht man eine Festsetzung darüber, dass ein bestimmter Ausdruck $A_1$ gleichbedeutend mit einem anderen Ausdruck $A_2$ sein soll, wobei die Bedeutung des anderen Ausdrucks $A_2$ als bekannt vorausgesetzt wird und $A_1$ die Bedeutung annehmen soll, die $A_2$ hat"[1].

Ueber die Realität wird mittels einer derartigen Definition nichts ausgesagt; eine Nominaldefinition ist eine blosse Konvention über die Verwendung sprachlicher Ausdrücke. Vor allem in bezug auf die Beantwortung der für die Entscheidungstheorie natürlich zentralen Frage (2) sagt eine Nominaldefinition nichts aus.

So wird die Wahrscheinlichkeit bisweilen definiert als Gewichtungsfaktor, der verschiedenen möglichen Ereignissen zugeordnet werden kann[2]. Dadurch gelingt aber keine Deutung des Wahrscheinlichkeitsbegriffs, denn das Problem verschiebt sich einfach auf die Interpretation des Terminus Gewichtungsfaktor.

Einen etwas anderen Charakter haben die beiden folgenden (Nominal-)Definitionen. Menges versteht unter Wahrscheinlichkeit ein

---

[1] Opp, S. 93

[2] Vgl. Schlaifer, S. 16

"Mass für die Neigung eines möglichen Ereignisses, sich realiter durchzusetzen, zum Vorschein zu kommen, verwirklichtes Ereignis zu werden"[1].

Dieser Umschreibung entgegengesetzt ist die Definition der Wahrscheinlichkeit als "Mass für unsere Unwissenheit"[2]. Beide Definitionen enthalten eine metaphysische Komponente[3], was an sich nicht erstaunlich ist, denn eines der ältesten philosophischen Probleme - das Determinismus/Indeterminismusproblem - ist eng mit dem Wahrscheinlichkeitsbegriff verknüpft.

Nach dem metaphysischen Determinismus ist in der Natur nichts zufällig, wohl aber unser Wissen über die Natur unvollständig. Diese Lehre endet in einem probabilistischen Subjektivismus, der in der zweiten genannten Definition im Ausdruck "Unwissenheit" offenbar wird[4].

Als Variante dieser subjektivistischen Definition lässt sich die häufig vertretene Deutung der Wahrscheinlichkeit als Ueberzeugungsgrad oder Grad des Glaubens auffassen[5]. Sie wird hauptsächlich von Entscheidungstheoretikern vertreten. Dabei erstaunt es nicht, dass eine Definition aus dieser Sicht "aktorbezogen" ausfällt.

Die Definition von Menges lässt sich demgegenüber der Indeterminismus-Position zurechnen, die davon ausgeht, dass die Natur selbst stochastische Prozesse hervorbringt. Es soll hier nicht mehr weiter auf diese Kontro-

---

[1] Menges (2), S. 21

[2] Stegmüller (3), S. 65

[3] Unter Metaphysik wird die Lehre vom Seienden verstanden. Eine metaphysische Aussage über die Struktur der Welt kann nicht anhand der Realität überprüft werden. Der Satz "alles in der Welt ist determiniert" ist beispielsweise eine solche Aussage, vgl. Stegmüller (1), S. 180

[4] Vgl. Stegmüller (3), S. 65 ff.

[5] Vgl. Ewart/Ford/Lin, S. 4; Peters/Summers, S. 6; Winkler, S. 16

verse eingegangen werden, hingegen sei noch vermerkt, dass die beiden Auffassungen nicht unbedingt monistisch vertreten werden müssen, sondern dass auch dualistische Konzepte existieren[1].

Wenn im folgenden die drei "klassischen" Wahrscheinlichkeitsbegriffe behandelt werden, wird man die erörterten Nominaldefinitionen nicht ganz ausser Acht lassen können, denn die drei Begriffstypen lassen sich als operationale Definitionen der Nominaldefinitionen auffassen.

b) Die drei klassischen Wahrscheinlichkeitskonzepte

ba) Die logische Wahrscheinlichkeit

Logische Wahrscheinlichkeit wird definiert als Anzahl der günstigen Fälle dividiert durch die Anzahl der möglichen Fälle. Man nennt diese Wahrscheinlichkeit auch Apriori-Wahrscheinlichkeit, mathematische, rationale, deduktive oder Merkmalswahrscheinlichkeit, spricht von der klassischen Wahrscheinlichkeitsdefinition oder Gleichmöglichkeitsdefinition[2].

"Der entscheidende Punkt in dieser klassischen Definition besteht ... darin, dass die Berechnung der Wahrscheinlichkeiten auf die Aufzählung von Möglichkeiten zurückgeführt wird"[3].

Wenn kein zureichender Grund für die Annahme besteht, dass sich eine Möglichkeit leichter verwirklichen werde als eine andere, betrachtet man alle Möglichkeiten als gleichwahrscheinlich (Prinzip des unzureichenden Grundes oder

[1] Vgl. Stegmüller, S. 111

[2] Vgl. Menges (2), S. 25; Stegmüller (4), S. 252

[3] Stegmüller (4), S. 252

Indifferenzprinzip[1]). Diese Symmetriebedingungen sind allerdings nur in den seltensten Fällen gegeben[2].

Vertreten wurde dieser Begriff von den Klassikern der Wahrscheinlichkeitstheorie, J. Bernoulli und P.-S. de Laplace.

bb) Die objektive Wahrscheinlichkeit

Die objektive Wahrscheinlichkeit (auch Häufigkeits-, Aposteriori-, frequentistische, statistische, empirische, induktive oder Massenwahrscheinlichkeit genannt)

"bezieht sich auf die tatsächlichen Realisationen ... eines Merkmals eines stochastischen Vorganges"[3].

Stegmüller unterscheidet zwei Varianten der "Häufigkeitstheorie". In der historisch ursprünglicheren Variante wird der Begriff der Wahrscheinlichkeit als Grenzwert von relativen Häufigkeiten definiert; sie wird deshalb "Limestheorie der statistischen Wahrscheinlichkeit" genannt[4].

An die Stelle der Limestheorie tritt nach der zweiten Variante die sogenannte "Propensitiv-Theorie". Diese verzichtet auf eine definitorische Rückführung des Begriffs der Wahrscheinlichkeit auf bereits verfügbare Begriffe (Reduktionismus) und führt statt dessen den Begriff als "theoretische Disposition" ein. Wahrscheinlichkeiten werden nicht mehr als Merkmale von Ereignisfolgen gedeutet, sondern als

[1] Menges (2), S. 25; Stegmüller (2), S. 66 f.

[2] Stegmüller (4), S. 252

[3] Menges (2), S. 25

[4] Stegmüller (4), S. 29

"Eigenschaften der experimentellen Anordnungen ..., welche derartige Folgen hervorrufen"[1].

Hauptvertreter der Limestheorie sind R. von Mises, H. Reichenbach und R.A. Fisher[2]; die Propensity-Theorie wird vertreten von K.R. Popper, R.N. Giere und P. Suppes[3].

bc) Die subjektive Wahrscheinlichkeit

Recht heterogen sind die Interpretationen des Begriffs der subjektiven (auch personellen, induktiven) Wahrscheinlichkeit. Ausgangspunkt ist dabei fast immer die genannte (Nominal-)Definition der Wahrscheinlichkeit als Ueberzeugungsgrad. Menges unterscheidet innerhalb des Subjektivismus zwei Richtungen: die intuitionistische oder introspektive (B. de Finetti, B.O. Koopmann, E. Halphen) und die bayesianische (F.P. Ramsey, L.J. Savage)[4]. Beide Richtungen definieren die subjektive Wahrscheinlichkeit operational als

"maximaler Wettquotient (zu dem X auf E zu wetten bereit wäre)"[5].

Unabhängig voneinander haben Ramsey und de Finetti gezeigt, dass bei einem kohärenten Wettsystem alle Axiome der mathematischen Wahrscheinlichkeitstheorie erfüllt werden[6]. Hingegen gehen die Meinungen, worüber man diese

---

[1] Stegmüller (4), S. 245

[2] Vgl. Carnap, S. 195

[3] Vgl. Stegmüller (4), S. 245 ff.

[4] Menges (1), S. 33 f.

[5] Stegmüller (4), S. 226

[6] Stegmüller (2), S. 81 f.

Wetten abschliessen kann, auseinander. Nach Savage kann es auch sinnvolle Wetten auf Allsätze geben, de Finetti lehnt diese Auffassung ab[1].

Viele Vertreter der bayesianischen Richtung, unter ihnen vor allem H. Raiffa und R. Schlaifer, vertreten den Standpunkt, dass alle Wahrscheinlichkeiten notwendigerweise subjektiv sind. Mit dem Terminus "subjektiv" wollen sie zum Ausdruck bringen, dass verschiedene Individuen dem gleichen Ereignis unterschiedliche Wahrscheinlichkeiten zuordnen, weil die Informationsbasis nicht dieselbe ist. Gleiche Information würde zu identischen Wahrscheinlichkeitsbeurteilungen führen[2]. Historischen relativen Häufigkeiten wird bei der Bestimmung der subjektiven Wahrscheinlichkeit grosser Wert beigemessen, wo immer sie zur Verfügung stehen, sind sie - nebst anderen Erfahrungen - zu verwenden[3].

Andererseits gibt es auch Bayesianer, die davon ausgehen, dass zwei Arten von Wahrscheinlichkeiten existieren, die frequentistische und die subjektive. Sie sind aber der Ansicht, dass in den meisten realen Situationen kein Wissen über relative Häufigkeiten verfügbar ist, so dass subjektive Ueberzeugungsgrade von grösster Bedeutung (insbesondere für die Entscheidungstheorie) sind[4]. Für sie sind subjektive Wahrscheinlichkeiten Schätzungen oder Werte, die aufgrund von nicht dem Häufigkeitstyp zuzuordnenden Erfahrungen zustandekommen.

---

[1] Stegmüller (4), S. 232

[2] Vgl. Hadley, S. 12; Schlaifer, S. 206

[3] Peters/Summers, S. 6 f.; Raiffa/Schlaifer, S. 5; Schlaifer, S. 207

[4] Vgl. Bühlmann/Loeffel/Nievergelt, S. 14; Thompson, S. 23

Für die starke Verbreitung des Verständnisses der Wahrscheinlichkeit als subjektive war vor allem eine permanente Schwierigkeit der Limestheorie bei der Interpretation der Wahrscheinlichkeit von Einzelereignissen verantwortlich. Dieser Punkt wird bei der Kritik der einzelnen Konzpete näher erörtert werden; hier sei nur vermerkt, dass dieses Problem in einigen subjektiven Interpretationen des Begriffes "Wahrscheinlichkeit" seinen Niederschlag findet[1].

Ein verbindendes Element aller subjektiver Auffassungen ist das Prinzip des Lernens aus der Erfahrung: Wahrscheinlichkeitsverteilungen können aufgrund zusätzlicher Information beliebig oft revidiert werden. Mit wachsender Information konvergieren die dadurch gewonnenen Aposteriori-Wahrscheinlichkeiten unabhängig von den urspünglichen Apriori-Werten gegen "ein und dieselbe interpersonelle Wahrscheinlichkeit"[2]. Die Revisison geschieht mit Hilfe des Bayes-Theorems.

Vertreter einer besonderen Variante der subjektiven Interpretation sind Carnap und Wittgenstein, die ihre Ueberlegungen auf dem klassischen Indifferenzprinzip aufbauen[3]. Eine Besonderheit der Theorien dieser beiden Philosophen ist, dass sie sich nicht wie die bisher behandelten Theorien auf Mengen und Ereignisse beziehen, sondern auf Sätze[4]. Merkwürdigerweise hat ein vehementer Kämpfer gegen

[1] Vgl. Peters/Summers, S. 6; Schlaifer, S. 206 f.; Winkler, S. 16

[2] Stegmüller (4), S. 232

[3] Vgl. Carnap, S. 197 ff.; Wittgenstein, S. 64 f.

[4] Stegmüller spricht von einer "linguistisch formulierten Theorie", Stegmüller (2), S. 77
Weingartner von einer "sematischen oder syntaktischen Interpretation", Weingartner, S. 220

den Subjektivismus, Popper, ebenfalls einen derartigen Wahrscheinlichkeitsbegriff eingeführt und spricht in diesem Zusammenhang von "logischer Wahrscheinlichkeit"[1]. Auch Carnap betont, dass es sich bei der induktiven Wahrscheinlichkeit um einen Begriff von "rein logischer Natur"[2] handle. Er grenzt ihn damit sowohl gegenüber dem statistischen Wahrscheinlichkeitsbegriff, der eine Tatsache behaupte[3], wie auch gegenüber einer psychologischen Interpretation im Rahmen des Subjektivismus ab[4]. Die semantische Interpretation wird dem subjektiven Wahrscheinlichkeitsbegriff zugeordnet, weil auch sie auf die Bedeutung des Grundsatzes des Lernens auf der Erfahrung verweist und aus diesem Grunde das klassische Indifferenzprinzip modifiziert[5].

c) Monistische und dualistische Auffassungen

Sowohl die objektive, (vor allem die auf der Limestheorie aufbauende) wie auch die subjektive Position werden häufig monistisch aufgefasst, d.h., dass die Vertreter einer Richtung die andere Konzeption ablehnen und bekämpfen. Es wurde aber bereits darauf hingewiesen, dass auch beide Positionen miteinander oder besser nebeneinander vertreten werden[6], in dem Sinne, dass von der Existenz zweier verschiedener, voneinander unabhängiger Wahrscheinlichkeitsbegriffe ausgegangen wird[7].

[1] Popper, S. 81 ff.

[2] Carnap, S. 195

[3] A.a.O.

[4] Stegmüller (2), S. 59

[5] Vgl. Carnap, S. 202 ff.; Stegmüller (2), S. 66 ff.

[6] Vgl. S. 37

[7] Vgl. Carnap, S. 198; Stegmüller (3), S. 111

d) Kritik der einzelnen Konzeptionen

Für den logischen Wahrscheinlichkeitsbegriff wurde bereits vermerkt, dass die zugrundeliegenden Symmetriebedingungen nur in wenigen Fällen gegeben sind, so dass er hier ausser Acht gelassen werden kann.

Stegmüller hat sämtliche je gegen die Limestheorie eingewendeten Argumente sehr ausführlich dargestellt und gewürdigt[1]. Er kommt zum Schluss, dass die Limestheorie als logisch widerlegt gelten kann. Als Einwand wird in der Stegmüllerschen Argumentensammlung auch die Schwierigkeit der Häufigkeitstheoretiker bei der Interpretation des Sprechens über die Wahrscheinlichkeit von Einzelereignissen, (die bereits erwähnt wurde) genannt[2]. Weil das Einzelereignis in der Optik der Entscheidungstheorie von grosser Bedeutung ist, soll hier noch kurz auf dieses Argument eingegangen werden.

Der mathematische Statistiker geht normalerweise von der folgenden "Fundamentalregel" aus:

"Statistische Aussagen beziehen sich nie auf ein Einzelereignis ..., sondern nur auf Gesamtheiten vieler Ereignisse ... Die Anwendung statistischer Resultate auf einen Einzelfall ist unzulässig"[3].

Allerdings bleibt selbst ein Klassiker der Limestheorie wie H. Reichenbach nicht bei einer so einschränkenden Formulierung stehen, sondern sucht nach einem

"Ausweg, der trotz Festhaltung der Häufigkeitsdeutung zu einer befriedigenden Lösung für das Problem der Einzel-

[1] Stegmüller (4), S. 32 ff.

[2] Stegmüller (4), S. 33 und 35 f.

[3] Hengst, S. 46

aussagen führt"[1].

Seine Konstruktion, die schliesslich zum Begriff der Beurteilung, "anstelle des unhaltbaren Begriffs der Wahrscheinlichkeit einer Einzelaussage" führt, mutet allerdings reichlich gewunden an[2]. Das Resultat dieser Bemühungen - den Begriff "Beurteilung" - könnte man schliesslich wohl auch als subjektive Wahrscheinlichkeit bezeichnen.

Auch Popper führt für die Wahrscheinlichkeit eines Einzelereignisses einen neuen Begriff ein: den der "formalistischen Wahrscheinlichkeit" und meint, dass dieser Begriff eine "Brücke zur subjektiven Theorie" schlage. Er erklärt sich auch damit einverstanden, diesen Begriff als "Grad des vernunftgemässen Wissens" zu interpretieren[3].

Dass selbst die klassische mathematische Statistik, die in der Regel ohne nähere Begründung auf der Häufigkeitsdeutung aufbaut, um eine letztlich subjektive Interpretation nicht herumkommt, zeigt die Verwendung der Begriffe des Konfidenzintervalls und des Konfidenzkoeffizienten, die geschaffen wurde, um eine Unterscheidung zum "legitimen Wahrscheinlichkeitsbegriff"[4] zu ermöglichen. Auch diese beiden Begriffe können ohne weiteres als Ueberzeugungsgrade gedeutet werden.

Diese Schwierigkeiten der Limestheorie bei der Deutung von Aussagen über Einzelereignisse widerlegen den objektiven Wahrscheinlichkeitsbegriff nicht, sie weisen aber

[1] Reichenbach, S. 182

[2] Reichenbach, S. 183

[3] Popper, S. 163

[4] Yamane, S. 211

auf Lücken hin, die dieser Begriff offen lässt, was zumindest eine dualistische Konzeption als vorteilhaft erscheinen lässt. Nachdem daneben aber offenbar zwingende Argumente gegen die Limestheorie selbst vorliegen, scheint es, dass der monistische Subjektivismus eine Position darstellt, die sowohl der Limestheorie wie auch der dualistischen Variante überlegen sein könnte. Gegen den Subjektivismus wird von Seiten der Objektivisten allerdings eingewendet, dass diese Konzeption

"mit dem Bemühen, Einsicht in objektive Sachverhalte und Wirkungszusammenhänge zu gewinnen, d.h. mit dem Bestreben, die Realität aus zureichenden Gründen heraus zu erklären und zu begreifen"[1]

unverträglich sei.

Wird dieser Standpunkt von einem Limestheoretiker vertreten, so wird ihn der Subjektivist mit dem Hinweis auf die Unzulänglichkeit der Limestheorie, die ihre Vertreter nicht zur Kritik am Subjektivismus legitimiere, zurückweisen. Der Subjektivist könnte seinem Kritiker aber auch sein unterschiedliches Erkenntnisinteresse entgegenhalten: während der Objektivist, wie der zitierte Einwand zeigt, ein deskriptives Erkenntnisinteresse hat, ist der Subjektivist primär aus einem präskriptiven Interesse heraus an der Wahrscheinlichkeit interessiert[2].

Damit sind aber nicht alle Einwände gegen den Subjektivismus erledigt. Nach Stegmüller besteht das grosse Dilemma der subjektiven Theorie darin, dass aufgrund der subjektiven Interpretation nicht mehr zwischen derjenigen Form von Ungewissheit, die auf mangelnde Information zu-

[1] Menges (1), S. 34

[2] Vgl. den folgenden Abschnitt

rückzuführen ist und der Ungewissheit, die durch keine zusätzliche Information zu beseitigen ist, unterschieden werden kann. Die Quantenphysik habe aber gezeigt, dass die Natur selbst stochastische Prozesse hervorbringe[1]. Man könne allerdings mit diesem Einwand die subjektive Theorie nicht zwingend verwerfen, müsse aber bei ihrer Beibehaltung

"eine radikale Subjektivierung der Naturwissenschaften in Kauf nehmen"[2].

Wenn man diesen Standpunkt nicht akzeptieren kann, folgert Stegmüller, so bleibt nur die Möglichkeit, auf den Reduktionismus, der für die Limestheorie wie für den Subjektivismus kennzeichnend ist, zu verzichten. Unter "Reduktionismus" wird der Versuch verstanden, einen Begriff

"mittels bereits verfügbarer logischer und empirischer Begriffe zu definieren"[3].

Der Begriff Wahrscheinlichkeit wäre demgegenüber als "theoretische Disposition" einzuführen, was zu der Variante der Propensity-Theorie führen würde.

e) Wahrscheinlichkeit in der Entscheidungstheorie

Nachdem die hauptsächlichsten in der Literatur verwendeten Wahrscheinlichkeitskonzeptionen dargestellt und kurz gewürdigt wurden, gilt es nun, aus der Sicht der betriebswirtschaftlichen Entscheidungstheorie (als einer präskriptiven Theorie) Bilanz zu ziehen und sich auf einen zweckmässigen Begriff festzulegen.

In der Entscheidungstheorie werden Wahrscheinlichkeiten

[1] Stegmüller (3), S. 65 ff.

[2] Stegmüller (4), S. 250

[3] Stegmüller (4), S. 59 und 234 ff.; vgl. auch S. 35

möglichen Umweltzuständen zugeordnet. An anderer Stelle wurde darauf hingewiesen, dass man in den wenigsten Fällen davon ausgehen kann, dass sich eine gleiche Entscheidungssituation auch in Zukunft häufig wiederholen wird, sondern vielmehr einmalige, sich nicht wiederholende Entscheidungssituationen auftreten[1]. Zudem dürfte sehr oft auch das Wissen über historische relative Häufigkeiten beschränkt sein.

Das heisst aber nichts anderes, als dass der reine Objektivist (Limestheoretiker) in der betriebswirtschaftlichen Entscheidungstheorie praktisch nur Ungewissheitssituationen anerkennen würde, bei denen keine Wahrscheinlichkeiten für das Eintreten der relevanten Umweltzustände bekannt sind. Um überhaupt zu einem Entscheid zu kommen, müsste er eine der bekannten Entscheidungsregeln für Ungewissheitssituationen (rationale Indeterminiertheit) verwenden, deren Wahl aber willkürlich ist und nicht rational begründet werden kann[2].

Der Subjektivist wird dieses Vorgehen des Objektivisten als absurd bezeichnen, denn seiner Meinung nach ist es widersinnig, auf alle Informationen, die nicht vom Häufigkeitstyp sind, mit dem Hinweis auf das Objektivitätspostulat zu verzichten und stattdessen aufgrund einer willkürlich gewählten Entscheidungsregel zu entscheiden. Er wird im Gegenteil auf die Objektivität seines Verfahrens hinweisen, das durch die Gewinnung und Verarbeitung zusätzlicher Informationen, nach dem Prinzip des Lernens aus der Erfahrung, die zunächst von Aktor zu Aktor verschiedenen Apriori-Wahrscheinlichkeiten einander annähert.

---

[1] Vgl. S. 10

[2] Vgl. Bamberg/Coenenberg, S. 93 ff.

Es ist kaum von der Hand zu weisen, dass das subjektive Konzept dem Anspruch der Entscheidungstheorie, dem Aktor, der nach einer rationalen Begründung seiner Entscheide sucht, behilflich zu sein, weit mehr entspricht als das objektive (immer in der Variante der Limestheorie). Neben diesen pragmatischen Ueberlegungen sprechen zudem auch die logischen Einwände gegen die Limestheorie.

Es scheint, dass ein monistischer Subjektivismus, der Wahrscheinlichkeit als Ueberzeugungsgrad (nominal) definiert und alle verfügbaren Informationen (also auch historische relative Häufigkeiten) für die Beurteilung von Wahrscheinlichkeiten miteinbezieht, als die der Entscheidungstheorie adäquate Interpretation angesehen werden kann. Dabei setzt man sich dem Vorwurf Stegmüllers aus, man nehme eine radikale Subjektivierung der Naturwissenschaften in Kauf. Demgegenüber lässt sich aber einwenden, dass hier nur Wahrscheinlichkeiten, die ein Aktor möglichen Umweltzuständen zuschreibt, betrachtet werden und zwar geleitet durch präskriptive Interessen. In deren Sicht ist die Determinismus/Indeterminismusfrage überhaupt nicht von Bedeutung. Wahrscheinlichkeiten haben nur den Zweck, einen Entscheid des Aktors zu unterbauen, "rationaler" zu machen. U.E. ist durch die Festlegung auf den Subjektivismus keine metaphysische Position im Sinne des Determinismus indiziert, und es wird die Möglichkeit offengelassen, in anderen Bereichen andere Interpretationen (beispielsweise im Simme der Propensity-Theorie) zu verwenden.

### 1.31. Interpretation der Apriori-Verteilung

Vertreter einer dualistischen Wahrscheinlichkeitskonzeption - Limestheorie und subjektive Theorie - neigen im allgemeinen dazu, die Apriori-Wahrscheinlichkeiten dem

subjektiven Wahrscheinlichkeitsbegriff unterzuordnen[1]. Dabei wird bisweilen auch die Meinung vertreten, dass Apriori-Wahrscheinlichkeiten zwangsläufig subjektiv seien, denn eine Revision von objektiven Wahrscheinlichkeiten sei ein sinnloses Unterfangen[2].

Aus der Sicht der hier vertretenen Interpretation des Begriffes "Wahrscheinlichkeit"[3] sind Apriori-Wahrscheinlichkeiten jedenfalls subjektive Wahrscheinlichkeiten. Sie beinhalten alles verfügbare Wissen über die in Frage stehende Unsicherheit, wenn vorhanden auch historische relative Häufigkeiten. Dabei kann es durchaus auch sinnvoll sein, frequentistische Apriori-Wahrscheinlichkeiten aufgrund zusätzlicher Information zu revidieren: dann nämlich, wenn sich die Apriori-Verteilung auf ein Einzelereignis bezieht und nicht auf eine Gesamtheit von Ereignissen, wie die zugrundeliegende Häufigkeitsverteilung.

### 1.32. Die Apriori-Verteilung in der Bayes'schen Entscheidungstheorie

Die Grundprobleme der Bayes'schen Entscheidungstheorie lassen sich zwei Teilbereichen zuordnen[4]. Im einen Bereich stehen Probleme der Ermittlung einer Aposteriori-Verteilung bei gegebener Apriori-Verteilung im Vordergrund. Fragen der Herkunft und Begründung der Apriori-Verteilung bleiben ausgeklammert.

Die Probleme dieses Bereichs hangen entscheidend davon

[1] Ferschl, S. 56; Menges, S. 192

[2] Krautter, S. 103 ff.; Wehrli, S. 108

[3] Vgl. S. 46

[4] Lin/Schick, S. 173 f.

ab, ob es sich bei der Apriori-Verteilung um eine diskrete oder stetige Verteilung handelt. Die Ermittlung der Aposteriori-Verteilung für den diskreten Fall wurde in ihren Grundzügen im Kapitel "Einführung in das Bayes'sche Modell des Entscheidens", das sich,ohne explizit darauf hinzuweisen, auf diesen Fall beschränkt hat, skizziert. Es wurde gezeigt, dass - die Normalkonstellation[1] der Wahrscheinlichkeiten vorausgesetzt - die Revision der Apriori-Wahrscheinlichkeiten aufgrund der Beobachtungsresultate mit Hilfe des Bayes-Theorems geschieht. Ist die Apriori-Verteilung eine stetige Verteilung, so ändert sich an diesem Vorgehen grundsätzlich nichts. Das Bayes-Theorem für den stetigen Fall muss aber wie folgt geschrieben werden:

$$f''(s/r) = \frac{f'(s)\ f(r/s)}{\int_{-\infty}^{\infty} f'(s)\ f(r/s)\ ds} \qquad (2.8)$$

Die Anwendung der Gleichung (2.8) bietet keinerlei Schwierigkeiten, wenn es sich bei der Aprioridichte f'(s) und der Likelihoodfunktion f(r/s) um sehr einfache mathematische Funktionen handelt. Sobald diese Funktionen aber etwas komplexer werden, kann die Bestimmung des Integrals im Nenner von (2.8) zu einer rechentechnisch kaum mehr zu bewältigenden Aufgabe werden[2]. Die Versuche, diese Schwierigkeiten zu überwinden, haben schliesslich zum Konzept der "konjugierten Familien von Verteilungen" (conjugate family of distributions) geführt[3].

[1] Vgl. S. 18

[2] Winkler, S. 147

[3] Winkler, S. 147 ff.

befasst sich mit der Ermittlung von Apriori-Verteilungen oder mit anderen Worten: Mit der Quantifizierung von individuellen Beurteilungen unsicherer Ereignisse. Für diesen Zweck wurde eine ganze Reihe von Techniken entwickelt (Probability Encoding Techniques)[1].

Die vorliegende Arbeit ist dem ersten dieser Bereiche zuzuordnen.

## 1.33. Ausprägungen der Apriori-Verteilung

### 1.331. Diskrete Apriori-Verteilung

Die Apriori-Verteilung ist in diesem Falle eine Massefunktion einer diskret verteilten Zufallsvariablen, beispielsweise die in Abb. 4 dargestellte:

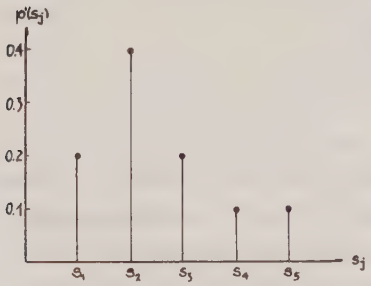


Abb. 4: Diskrete Apriori-Verteilung

### 1.332. Stetige Apriori-Verteilung

Die stetige Apriori-Verteilung ist eine Dichtefunktion, beispielsweise die in Abb. 5 dargestellte:

[1] Eine tabellarische Uebersicht über diese Methoden findet sich bei Brant, S. 167 und Spetzler/Staël von Holstein, S. 348

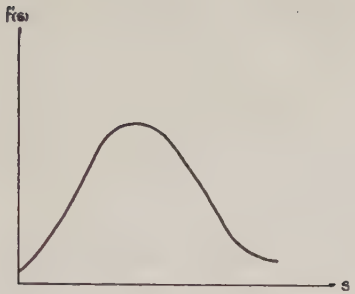


Abb. 5: Stetige Apriori-Verteilung

Die Anzahl der Zufallsvariablen, d.h. hier die Anzahl der möglichen Umweltzustände, ist in diesem Falle nicht begrenzt. Die möglichen Umweltzustände können allerdings in einer konkreten Entscheidungssituation zu einer endlichen Anzahl von Zustandsräumen zusammengefasst werden. Ein solcher Zustandsraum bezieht sich dann auf ein Intervall von Umweltzuständen, beispielsweise die Ausschussanteile $s_1$[0%, 1%], $s_2$[1%, 2%] usw.

Im folgenden wird zunächst eine einfache stetige Apriori-Verteilung - die Treppenverteilung - vorgestellt und anschliessend das Konzept der konjugierten Familien von Verteilungen behandelt.

1.332.1. Die Treppenverteilung

Bei der Treppenverteilung ist die Wahrscheinlichkeitsdichte f'(s) innerhalb eines Intervalls konstant. Bezogen auf ein Intervall spricht man etwa auch von "uniformer

Verteilung", "Gleich-" oder "Rechteckverteilung"[1].

Für den Fall von zwei Intervallen hat die Treppenverteilung folgendes Aussehen:

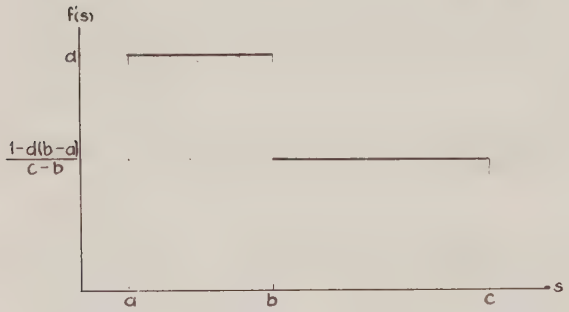


Abb. 6: Treppenverteilung über zwei Intervallen

Ihre Dichtefunktion ist:

$$f'(s) = \begin{cases} d & \text{für } a \le s < b \\ \frac{1-d(b-a)}{c-b} & \text{für } b \le s \le c \\ 0 & \text{sonst,} \end{cases}$$

mit $a<b<c$ und $(b-a)d \le 1$.

1.332.2. Konjugierte Familien von Verteilungen

Es wurde darauf hingewiesen, dass die Anwendung der Gleichung (2.8) dann Schwierigkeiten verursacht, wenn es

[1] Vgl. Sachs, S. 82; Winkler, S. 136 f.

sich bei der Apriori-Verteilung um eine komplexere Dichtefunktion handelt[1]. Das Konzept der konjugierten Familien von Verteilungen bietet eine Möglichkeit, diesen Schwierigkeiten zu begegnen, indem ein Mitglied einer solchen Familie als Apriori-Dichte verwendet wird. Natürlich hangen die zitierten Schwierigkeiten bei der Anwendung von (2.8) nicht allein von der Apriori-Dichte, sondern ebensosehr von der Likelihoodfunktion und damit vom verwendeten stochastischen Modell ab: Ist das in einem konkreten Entscheidfall als adäquat erachtete stochastische Modell bestimmt, so gilt es, eine Apriori-Dichte aus der Schar von Mitgliedern von der mit diesem stochastischen Modell konjugierten Familie von Verteilungen so auszuwählen, dass diese der Wahrscheinlichkeitsbeurteilung des Aktors am besten entspricht (fitting the conjugate distribution)[2]. Der Terminus "konjugiert" bezieht sich sowohl auf das Verhältnis zwischen Apriori- und Aposteriori-Verteilung, als auch auf die Verbindung zwischen Apriori- bzw. Aposteriori-Verteilung mit einem stochastischen Modell, bzw. mit der durch das stochastische Modell bestimmten Likelihoodfunktion.

Eine konjugierte Familie von Verteilungen hat folgenden Bedingungen zu genügen[3]:

- Sie soll mathematisch leicht zu handhaben sein;
- eine Familie von Apriori-Verteilungen soll eine Fülle von Formen aufweisen, damit ein breites Spektrum von Beurteilungen abgedeckt werden kann und
- sie sollte leicht zu interpretieren sein.

[1] Vgl. S. 47

[2] Vgl. Raiffa/Schlaifer, S. 58 ff. und Winkler, S. 158 f.

[3] Raiffa/Schlaifer, S. 43 f.; Winkler, S. 148 f.

Die erste dieser drei Bedingungen bezieht sich auf die Schwierigkeiten der Applikation des Bayes-Theorems auf den stetigen Fall. Mathematisch leicht zu handhaben ist eine Familie von Apriori-Verteilungen, wenn:

1. die Aposteriori-Verteilung bei gegebener Apriori-Verteilung und gegebenem stochastischen Modell ohne Schwierigkeiten hergeleitet werden kann;
2. die Aposteriori-Verteilung der gleichen Familie von Verteilungen angehört wie die Apriori-Verteilung, d.h. die beiden Verteilungen sich nur bezüglich der Werte ihrer Parameter unterscheiden und
3. die Ermittlung von Erwartungswerten möglich ist.

1.332.21. Binomial-Modell: Betaverteilung[1]

Die mit dem Binomial-Modell konjugierte Verteilung ist die Betaverteilung mit der Dichtefunktion

$$f'_{\beta}(\Theta;r',n') = \begin{cases} \frac{\Gamma(n')}{\Gamma(r')\,\Gamma(n'-r')}\,\Theta^{r'-1}(1-\Theta)^{n'-r'-1} & \text{für } 0\leq\Theta\leq 1 \\ 0 & \text{sonst,} \end{cases} \qquad (2.9)$$

mit $n'>r'>0$.

Die Funktion $\Gamma(y)$ heisst Gammafunktion und ist folgendermassen definiert[2]:

$$\Gamma(y) = \int_0^{\infty} x^{y-1}\,e^{-x}\,dx \qquad \text{für } y>0$$

[1] Vgl. Dück/Bliefernich, S. 328 ff.; Raiffa/Schlaifer, S. 53 und S. 216 ff.; Winkler, S. 149 ff.

[2] Vgl. Dück/Bliefernich, S. 325

Für ganzzahlige y ist

$$\Gamma(y) = (y-1)!$$

Falls n' und r' ganzzahlig sind, hat die Betaverteilung also folgende Form

$$f'_\beta(\theta;r',n') = \begin{cases} \frac{(n'-1)!}{(r'-1)!\,(n'-r'-1)!}\,\theta^{r'-1}\,(1-\theta)^{n'-r'-1} & \text{für } 0\leq\theta\leq 1 \\ 0 & \text{sonst.} \end{cases} \tag{2.10}$$

Korrekterweise müsste man, wenn θ im Intervall [0,1] liegt, von einer standardisierten Betaverteilung sprechen, der Einfachheit halber wird im folgenden aber dieser Zusatz weggelassen und unter Betaverteilung immer die definierte, standardisierte Verteilung verstanden.

Erwartungswert E und Varianz V der Betaverteilung sind

$$E(\theta;r,n) = \frac{r'}{n'} \tag{2.11}$$

$$V(\theta;r,n) = \frac{r'(n'-r')}{n'^2(n'+1)} \tag{2.12}$$

Die Betaverteilung weist einen sehr grossen Formenreichtum auf, erfüllt also das zweite an eine konjugierte Verteilung gestellte Postulat (vgl. Tab. 2).

| r' | n' | n'-r' | |
|---|---|---|---|
| $\frac{n}{2}$ | 2r | | Eingipflig symmetrisch |
| $<\frac{n}{2}$ | $>2r$ | | linkssteil |
| $>\frac{n}{2}$ | $<2r$ | | rechtssteil |
| $>1$ | | $>1$ | Eingipflig, häufigster Wert: $\frac{r-1}{n-2}$ |
| $\leq 1$ | | | Eingipflig, häufigster Wert 0 |
| | | $\leq 1$ | oder 1 oder U-förmig mit häufigsten Werten 0 und 1 |
| 1 | 2 | 1 | Uniforme Verteilung über dem ganzen Intervall |

Tab. 2: Formen von Betaverteilungen[1]

1.332.22. Poisson-Modell: Gammaverteilung[2]

Mit dem Poisson-Modell ist die Gamma-Verteilung konjugiert. Sie besitzt die Dichte

$$f'_{\gamma}(\lambda;r',t') = \begin{cases} \frac{e^{-\lambda t'}(\lambda t')^{r'-1}t'}{\Gamma(r')} & \text{für } \lambda \geq 0 \\ 0 & \text{sonst,} \end{cases} \qquad (2.13)$$

mit $r',t'>0$.

[1] Grafische Darstellungen von Betaverteilungen finden sich bei Raiffa/Schlaifer, S. 218 f.; Winkler, S. 151 f.

[2] Vgl. Dück/Bliefernich, S. 325 ff.; Raiffa/Schlaifer, S. 53 und S. 225 ff.; Winkler, S. 180

Ist r' ganzzahlig, so erhält man

$$f'_{\gamma}(\lambda;r',t') = \begin{cases} \dfrac{e^{-\lambda t'}(\lambda t)^{r'-1}t'}{(r'-1)!} & \text{für } \lambda \geq 0 \\ 0 & \text{sonst} \end{cases} \qquad (2.14)$$

Gammaverteilungen sind asymetrische Verteilungen mit positiver Schiefe[1] und dem Erwartungswert

$$E(\lambda;r',t') = \frac{r'}{t'} \qquad (2.15)$$

sowie der Varianz

$$V(\lambda;r',t') = \frac{r'}{t'^2} \qquad (2.16)$$

1.332.23 Normal-Modell[2]

Unter dem Normal-Modell soll in diesem Zusammenhang die Annahme verstanden werden, dass eine Grundgesamtheit oder der Output eines Prozesses $(\mu,\sigma)$ normalverteilt sei. Die Likelihoodfunktion und die damit verbundene konjugierte Apriori-Verteilung hangen davon ab, welche Paramenter als Umweltzustände in ein Entscheidmodell eingehen und ob gegebenenfalls einer dieser Parameter als bekannt vorausgesetzt werden kann. Dabei geht natürlich derjenige Parameter in das Entscheidmodell ein, von dem man annimmt,

[1] Eine grafische Darstellung findet sich bei Raiffa/Schlaifer, S. 226

[2] Vgl. Raiffa/Schlaifer, S. 54 ff. und S. 290 ff.; Winkler, S. 162 ff.

dass er die Konsequenzen des Entscheides beeinflusse.

Folgende Fälle sind möglich:

1) Der Mittelwert $\mu$ geht als Umweltzustand in das Entscheidungsmodell ein, die Standardabweichung $\sigma$ ist bekannt.

2) $\sigma$ geht als Umweltzustand in das Entscheidungsmodell ein, $\mu$ ist bekannt.

3) Weder $\mu$ noch $\sigma$ sind bekannt. Als Umweltzustand geht $\mu$ in das Entscheidungsmodell ein.

4) Weder $\mu$ noch $\sigma$ sind bekannt, als Umweltzustand geht $\sigma$ in das Entscheidungsmodell ein.

5) Weder $\mu$ noch $\sigma$ sind bekannt und beide gehen als Umweltzustände in ein Entscheidungsmodell ein.

Ein Beispiel möge eine dieser Konstellationen verdeutlichen:

Ein Aktor habe zu entscheiden, ob für ein bestimmtes Produkt eine neue Verpackung eingeführt werden soll. Das betreffende Produkt wird in ca. 20'000 Detailhandelsgeschäften abgesetzt, wobei der durchschnittliche Jahresabsatz des Produktes all dieser Detailhandelsgeschäfte normalverteilt ist. Der Entscheid ist abhängig von der zu erwartenden Umsatzsteigerung oder, was auf dasselbe hinausläuft, von der zu erwartenden Erhöhung des Mittelwertes dieser Normalverteilung des Durchschnittabsatzes aller Geschäfte. Als Umweltzustand geht $\mu$ in das Entscheidmodell ein, und es kann wohl angenommen werden, dass die Standardabweichung $\sigma$ konstant bleibt, also als bekannt vorausgesetzt werden kann. Eine Stichprobe, die der Aktor gegebenenfalls zieht, um die Qualität seines Entscheides zu erhöhen,

soll dem Aktor Aufschluss geben über den neuen Mittelwert nach der Einführung der Verpackung. Es liegt hier also Fall 1 vor.

Im folgenden werden die Fälle 3, 4 und 5 nicht mehr weiter behandelt. Bei Fall 3 kann bei genügend grossem Stichprobenumfang ($n>50$) die Stichprobenvarianz, multipliziert mit dem Faktor $\frac{n}{n-1}$ als unverzerrter Schätzwert für $\sigma^2$ verwendet werden[1]. Wenn die genannte Bedingung erfüllt ist, unterscheidet sich Fall 3 nicht mehr von Fall 1. Fall 4 soll hier ausscheiden, weil er als wenig realistisch angesehen wird. Fall 5 schliesslich wird hier weggelassen, weil er mit einer noch zu formulierenden Bedingung - dass als einzige stetige Konsequenzenfunktionen lineare Konsequenzenfunktionen verwendet werden[2] - im Widerspruch steht. Aber auch in der Literatur ist diese Konstellation nicht unumstritten, weil die Bestimmung der Konsequenzen in einem solchen Falle grosse Schwierigkeiten mit sich bringt[3].

a) $\sigma$ bekannt, $\mu$ unbekannt: Normalverteilung

Wenn $\mu$ als Umweltzustand in das Entscheidungsmodell eingeht, so ist die mit dem Normal-Modell konjugierte Apriori-Verteilung eine Normalverteilung. Parameter dieser Apriori-Verteilung sind der Erwartungswert $m'$ und die Standardabweichung $\sigma'$.

Aus später ersichtlich werdenden Gründen ist es üblich, einen neuen Parameter

$$n' = \frac{\sigma^2}{\sigma'^2}$$

---

[1] Yamane, S. 648 f.

[2] Vgl. S. 60

[3] Winkler, S. 182

einzuführen. Mit dieser Notation ist die konjugierte Apriori-Verteilung für das Normal-Modell $\varphi(m', \frac{\sigma}{\sqrt{n'}})$ normalverteilt.

b) μ bekannt, σ unbekannt: Gamma-2-Verteilung[1]

Aus mathematischen Gründen geht in diesem Falle nicht σ als Umweltzustand in das Entscheidmodell ein, sondern $h = \frac{1}{\sigma^2}$[2]. Die mit dem Normal-Modell verbundene Apriori-Verteilung ist dann eine Gamma-2-Verteilung mit der Dichte

$$f'_{\gamma 2}(h;\nu',v') = \begin{cases} \dfrac{e^{-\frac{1}{2}h\nu'v'} \; (\frac{1}{2}h\nu'v')^{\frac{1}{2}\nu'-1} \; (\frac{1}{2}\nu'v')}{\Gamma(\frac{1}{2}\nu')} & \text{für } h \geq 0 \\ 0 & \text{sonst,} \end{cases} \tag{2.17}$$

wenn $\nu', v' > 0$.

Ist der Ausdruck $(\frac{1}{2}\nu')$ ganzzahlig, so kann die Gammafunktion im Nenner durch $(\frac{1}{2}\nu'-1)!$ ersetzt werden

Erwartungswert und Varianz der Gamma-2-Verteilung sind

$$E(h;\nu',v') = \frac{1}{v'} \; , \tag{2.18}$$

$$V(h;\nu',v') = \frac{1}{\frac{1}{2}\nu' v'^2} \; . \tag{2.19}$$

---

[1] Raiffa/Schlaifer, S. 226 f.

[2] Raiffa/Schlaifer nennen h "precision of the process", S. 290

## 1.4 Die Art der Konsequenzenfunktion

Es ist üblich, den Zusammenhang zwischen Aktionen, Umweltzuständen und Konsequenzen durch die Funktion

$$k_{ij} = f(a_i, s_j)$$

auszudrücken. Dargestellt wird dieser Zusammenhang - wie in Teil I dieser Arbeit gezeigt - in der Regel mittels einer Konsequenzentabelle oder eines Entscheidungsbaumes.

Dabei wird allerdings von der impliziten Annahme ausgegangen, die Konsequenzenfunktion einer Aktion $a_i$

$$k(a_i) = f(\bar{a}_i, s_j)$$

sei eine diskrete Funktion (eine Tabelle müsste ja sonst unendlich viele Spalten, bzw. ein Entscheidungsbaum unendlich viele Aeste aufweisen) oder mit anderen Worten, es sei in einem Entscheidungsmodell nur eine begrenzte, endliche Anzahl von Situationsvariablen vorhanden.

Es ist aber durchaus möglich und in manchen Fällen auch realistisch, die Zahl der Umweltzustände nicht begrenzen und demzufolge mit einer stetigen Konsequenzenfunktion zu arbeiten[1].

Eine stetige Konsequenzenfunktion hat eine stetige Apriori-Verteilung zur Folge, die umgekehrte Implikation gilt aber nicht, denn, wie im Abschnitt über die stetige Apriori-Verteilung erwähnt, können in Zustandsräumen Intervalle von Umweltzuständen zusammengefasst werden[2]. Obschon in einem solchen Falle den Intervallen je ein einziger, dis-

[1] Vgl. Winkler, S. 348 f.

[2] Vgl. S. 49

kreter Konsequenzenwert zugeordnet wird, kann die Apriori-Verteilung durchaus eine stetige sein. Wenn es sich nicht um eine sehr einfache Apriori-Verteilung handelt, sind aber Entscheidungen mit stetiger Apriori-Verteilung und diskreter Konsequenzenfunktion rechentechnisch kaum mehr zu bewältigen. Aus diesem Grunde werden hier nur diskrete Apriori-Verteilungen mit diskreten Konsequenzenfunktionen und stetige Apriori-Verteilungen mit stetigen Konsequenzenfunktionen kombiniert.

Eine stetige Konsequenzenfunktion kann natürlich beliebig viele Formen annehmen. Im Rahmen dieser Arbeit soll nur eine - die lineare Konsequenzenfunktion - behandelt werden.

## 2. Kombination der typologischen Merkmale

Durch Kombination der Ausprägungen der vier behandelten typologischen Merkmale ergeben sich jetzt die Typen von 2-Aktionen-Problemen, die hier behandelt werden sollen. Als Hilfsmittel für eine übersichtliche Darstellung dieser Typenbildung sei die morphologische Methode[1] verwendet.

[1] Vgl. Jantsch, S. 175 ff.

In Tab. 3 sind die Merkmale mit ihren Ausprägungen aufgelistet:

| Merkmale | Ausprägungen | | | |
|---|---|---|---|---|
| $m_{1x}$ Zusätzliche Informationen | $m_{11}$ keine | $m_{12}$ mit gegebenem Umfang | $m_{13}$ mit variablem Umfang | $m_{14}$ mit variablem Umfang, sequentiell |
| $m_{2x}$ Stochastisches Modell | $m_{21}$ Bernoulli | $m_{22}$ Hypergeometrisches | $m_{23}$ Poisson | $m_{24}$ Normal |
| $m_{3x}$ Apriori-Verteilung | $m_{31}$ diskrete | $m_{32}$ stetige | | |
| $m_{4x}$ Konsequenzenfunktion | $m_{41}$ diskrete | $m_{42}$ stetige | | |

Tab. 3: Die typologischen Merkmale und ihre Ausprägungen

Theoretisch sind 64 Kombinationen von Merkmalsausprägungen möglich. Diese Anzahl wird durch die folgenden Restriktionen reduziert:

- Restriktion 1: Wenn keine zusätzliche Information erhoben wird, ist das stochastische Modell überflüssig: es fallen alle Verknüpfungen $m_{11}$-$m_{2x}$ weg.
- Restriktion 2: Die Verknüpfungen $m_{31}$-$m_{42}$ fallen weg, weil sich diskrete Apriori-Verteilungen und stetige Konsequenzenfunktion ausschliessen[1].
- Restriktion 3: Die Verknüpfungen $m_{32}$-$m_{41}$ werden aus rechentechnischen Gründen nicht miteinbezogen[2].

[1] Vgl. S. 60

[2] Vgl. a.a.O.

Im folgenden Teil werden der Reihe nach drei Pakete von Typen behandelt:

1)
$$m_{11} - m_{31} - m_{41}$$
$$m_{11} - m_{32} - m_{42}$$

2)
$$m_{12} - m_{2x} - m_{31} - m_{41}$$
$$m_{12} - m_{2x} - m_{32} - m_{42}$$

3)
$$m_{13} - m_{2x} - m_{31} - m_{41}$$
$$m_{13} - m_{2x} - m_{32} - m_{42}$$

Ein kurzer Schlussabschnitt befasst sich schliesslich noch mit dem sequentiellen Informationsentscheid.

# TEIL III DIE ENTSCHEIDUNG

## 1. Die Entscheidung ohne zusätzliche Informationsaktivität

Das in der Literatur praktisch unbestrittene Entscheidungskriterium für Risikosituationen, "das rationale Entscheidungskriterium schlechthin"[1], ist der Nutzenerwartungswert. Dieses Kriterium hat eine ganze Reihe von Namen: auf eine Nutzenfunktion angewandt (oder bei Risikoneutralität) sind die Begriffe "Bayes-Regel" bzw. "-Kriterium", "μ-Kriterium" und "Bernoulli-Kriterium" synonym[2].

Nach diesem Kriterium und unter den hier getroffenen Voraussetzungen der Linearität der Nutzenfunktion, ist diejenige Aktion a* optimal, die die maximalen erwarteten Konsequenzen $Ek(a_i)$ bzw. die minimalen erwarteten Opportunitätskosten $Eo(a_i)$ sämtlicher zur Verfügung stehender Aktionen aufweist:

$$Ek(a^*) = \max_i Ek(a_i), \tag{3.1}$$

$$Eo(a^*) = \min_i Eo(a_i). \tag{3.2}$$

Bei Vorliegen einer diskreten Apriori-Verteilung und einer diskreten Konsequenzenfunktion gilt

$$Ek(a_i) = \sum_{j=1}^{J} p'(s_j)\, k_{ij} \tag{3.3}$$

und

$$Eo(a_i) = \sum_{j=1}^{J} p'(s_j)\, o_{ij}. \tag{3.4}$$

[1] Menges (2), S. 191

[2] Bamberg/Coenenberg, S. 78

Bei einer stetigen Apriori-Verteilung mit dem Erwartungswert E(S') und der stetigen, linearen Konsequenzenfunktion

$$k(a_i) = m_i + t_i s,$$

ist der Konsequenzenerwartungswert einer Aktion

$$Ek(a_i) = E(m_i + t_i S') = m_i + t_i E(S') \quad [1]. \tag{3.5}$$

Es ist ersichtlich, dass in diesem Falle die Kenntnis des Erwartungswertes der Apriori-Verteilung ausreicht, um zu einem Entscheid zu gelangen.

Es seien

$$k(a_1) = m_1 + t_1 s$$

und

$$k(a_2) = m_2 + t_2 s$$

die Konsequenzenfunktionen eines 2-Aktionen-Problems. Der "Break-even-Punkt", bei dem die Konsequenzenerwartungswerte der beiden Aktionen gerade gleich gross sind, kann wie folgt ermittelt werden:

$$Ek(a_1) = Ek(a_2)$$

$$m_1 + t_1 E(S') = m_2 + t_2 E(S')$$

Die beiden Aktionen sind gleichwertig, wenn die Gleichung

$$\frac{m_1 - m_2}{t_2 - t_1} = E(S'), \quad t_1 \neq t_2$$

erfüllt ist.

[1] Vgl. hier und im folgenden Winkler, S. 348 ff.

$s_b$ sei der Break-even-Punkt

$$s_b = \frac{m_1 - m_2}{t_2 - t_1} . \tag{3.6}$$

Damit ein Break-even-Punkt existiert, müssen sich die beiden Konsequenzengeraden im ersten Quadranten schneiden. Wenn dies nicht der Fall ist, liegt eine "natürliche Ordnung" vor, d.h., dass die Konsequenzen der einen Aktion bezüglich jeden Umweltzustandes vorteilhafter sind als diejenigen der anderen Aktion[1]. Die optimale Aktion kann in einem solchen Falle allein aufgrund der Konsequenzenverläufe gewählt werden, die Apriori-Verteilung spielt dabei keine Rolle.

Ein Break-even-Punkt existiert, wenn eine der beiden folgenden Konstellationen gegeben ist:

I. $m_1 > m_2$

und $t_1 < t_2$

oder

II. $m_1 < m_2$

$t_1 > t_2$.

Wie eben gezeigt, gilt

$$Ek(a_1) = Ek(a_2)$$

wenn $E(S') = s_b$.

[1] Zum Begriff der natürlichen Ordnung vgl. Müller, S. 65

In diesem Falle ist

$$m_1 + t_1 E(S') = m_2 + t_2 E(S').$$

Wenn Konstellation I gegeben ist und $E(S') > s_b$, gilt

$$m_1 + t_1 E(S') < m_2 + t_2 E(S')$$

bzw. $$Ek(a_1) < Ek(a_2)$$

und analog $$Ek(a_1) > Ek(a_2),$$

wenn $$E(S') < s_b.$$

Den folgenden Ueberlegungen liegt jeweils diese Konstellation I zugrunde (vgl. Abb. 7).

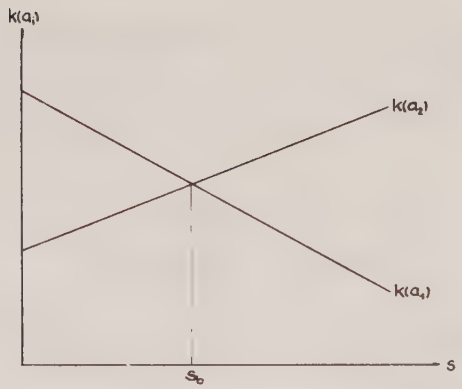


Abb. 7: Konsequenzenfunktionen und Break-even-Punkt eines 2-Aktionen-Problems

Die Opportunitätskostenfunktionen dieses 2-Aktionen-Problems sind

$$o(a_1) = \begin{cases} 0 & \text{für } s \leq s_b \\ (m_2 - m_1) + (t_2 - t_1)s & \text{für } s > s_b \end{cases} \tag{3.7}$$

$$o(a_2) = \begin{cases} (m_1 - m_2) + (t_1 - t_2)s & \text{für } s < s_b \\ 0 & \text{für } s \geq s_b \end{cases} \tag{3.8}$$

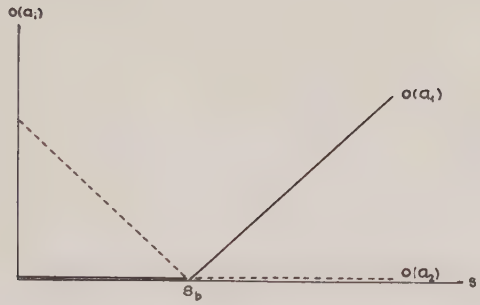


Abb. 8: Opportunitätskostenfunktionen eines 2-Aktionen-Problems

Die entsprechenden Opportunitätskosten-Erwartungswerte sind

$$Eo(a_1) = \int_{s_b}^{\infty} [(m_2-m_1)+(t_2-t_1)s]\, f'(s)\, ds \tag{3.9}$$

und

$$Eo(a_2) = \int_{-\infty}^{s_b} [(m_1-m_2)+(t_1-t_2)s]\, f'(s)\, ds, \tag{3.10}$$

wobei f'(s) die Aprioridichte symbolisiert.

Mit Hilfe von Beziehung (3.6) können die Gleichungen (3.7) und (3.8) umgeformt werden:

$$o(a_1) = \begin{cases} 0 & \text{für } s \leq s_b \\ (t_2-t_1)(s - s_b) & \text{für } s > s_b \end{cases} \tag{3.11}$$

$$o(a_2) = \begin{cases} (t_2-t_1)(s_b - s) & \text{für } s < s_b \\ 0 & \text{für } s \geq s_b \end{cases} \tag{3.12}$$

Eingesetzt in (3.9) und (3.10) ergibt sich

$$Eo(a_1) = (t_2-t_1) \int_{s_b}^{\infty} (s - s_b)\, f'(s)\, ds \tag{3.13}$$

bzw.

$$Eo(a_2) = (t_2-t_1) \int_{-\infty}^{s_b} (s_b - s)\, f'(s)\, ds \; . \tag{3.14}$$

Nach diesen allgemeinen Ausführungen über den Entscheidungstyp stetige Apriori-Verteilung/stetige Konsequenzenfunktion soll im folgenden noch ein Spezialfall betrachtet werden.

Apriori-Verteilung sei eine Treppenverteilung mit zwei Intervallen; die beiden Intervalle werden durch den Break-even-punkt getrennt. Als Umweltzustand gehe eine Proportion $\Theta$, beispielsweise ein Auslastungsgrad, in das Entscheidungsmodell ein. Die beiden Intervalle werden nach unten und oben durch die Werte 0 und 1 begrenzt.

Die Dichtefunktion der Apriori-Verteilung ist

$$f'(\theta) = \begin{cases} \dfrac{1-c(1-s_b)}{s_b} & \text{für } 0 \le \theta \le s_b \\ c & \text{für } s_b \le \theta \le 1 \\ 0 & \text{sonst} \end{cases}$$

und hat folgendes Aussehen:

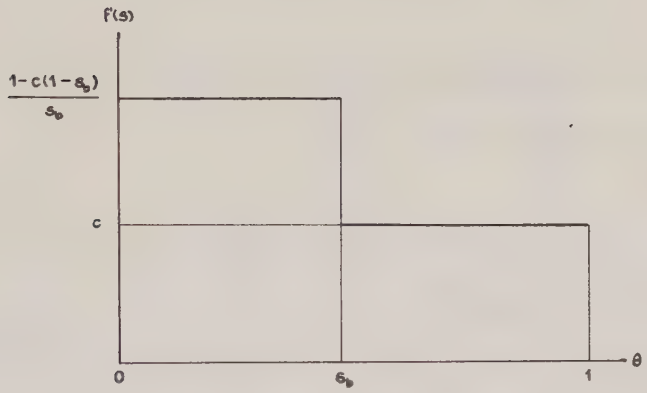


Abb. 9: Treppenverteilung mit den Intervallen $[0, s_b]$ und $[s_b, 1]$

Die Opportunitätskostenfunktion einer Aktion $a_1$ sei

$$o(a_1) = \begin{cases} a + b\theta & \text{für } s_b \le \theta \le 1 \\ 0 & \text{sonst} \end{cases}$$

Die erwarteten Opportunitätskosten dieser Aktion sind nach (3.13)

$$Eo(a_1) = b \int_{s_b}^{1} (\theta - s_b)\, f'(\theta)\, d\theta$$

$$= b \int_{s_b}^{1} (\theta - s_b)\, c\, d\theta$$

$$= b \left[ \left( \frac{c}{2} \theta^2 - c s_b \theta \right) \Big|_{s_b}^{1} \right].$$

Durch einige Umformungen erhält man

$$Eo(a_1) = \frac{-bs_b + b}{2} (c - cs_b)$$

oder

$$Eo(a_1) = \frac{-bs_b + b}{2} \int_{s_b}^{1} c\, d\theta.$$

Weil aber

$$a = -bs_b$$

erhält man schliesslich

$$Eo(a_1) = \frac{a + b}{2} \int_{s_b}^{1} c\, d\theta. \qquad (3.15)$$

Der Faktor $\frac{a+b}{2}$ drückt nun aber nichts anderes aus, als die Höhe der Opportunitätskosten in der Mitte des Intervalls $[s_b, 1]$ (vgl. Abb. 10). Das bestimmte Integral $\int_{s_b}^{1} c\, d\theta$ ist gleich der Wahrscheinlichkeit $p(s_b \leq \theta \leq 1)$.

Aus dem Beispiel lässt sich folgender verallgemeinernder Schluss ziehen: Bei einer stetigen, uniformen Apriori-Verteilung über einem Intervall, ist der Opportunitäts-

kosten-Erwartungswert einer Aktion - eine lineare Konsequenzenfunktion vorausgesetzt - gleich dem Produkt der Opportunitätskosten der Intervallmitte und dem bestimmten Integral mit den Grenzen dieses Intervalls.

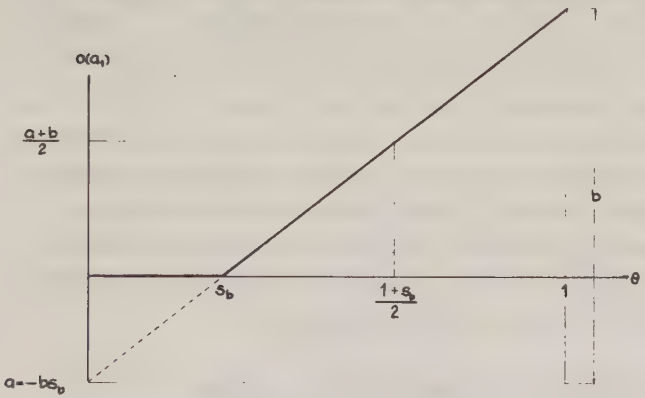


Abb. 10: Opportunitätskostenfunktion einer Aktion $a_1$

Die praktische Auswirkung dieser Erkenntnis ist, dass für die Ermittlung der erwarteten Opportunitätskosten in diesem Falle nicht die Gleichungen (3.9) und (3.10) bzw. (3.13) und (3.14) verwendet werden müssen, sondern die Anwendung der Beziehung (3.4) genügt, wenn die Werte für $p(a \leq s \leq b)$ bekannt sind.

## 2. Die Entscheidung mit zusätzlicher Informationsaktivität: gegebener Informationsumfang

Während im vorangehenden Abschnitt direkt nach der optimalen Aktion gefragt wurde, geht es in diesem Kapitel um das Auffinden einer optimalen Entscheidungsregel, die jedem möglichen Ergebnis einer Informationssuche die optimale Aktion zuschreibt.

Dieses Problem kann mit zwei grundsätzlich verschiedenen Methoden angegangen werden: mit der extensiven Methode einerseits oder mit der Normalform der Analyse anderseits. Diese Begriffe sind der Spieltheorie entlehnt, synonym werden auch die Termini "integrative Methode" (= Normalform) und "konstruktive Methode" (= extensive Methode) verwendet[1].

Erstaunlicherweise finden sich in der doch sehr umfangreichen Spezialliteratur über Bayes'sche Entscheidungstheorie nur recht selten Hinweise über die Existenz dieser beiden Methoden. In der Regel wird ausschliesslich eine der beiden Methoden angewandt, in jüngerer Zeit ist dies vorwiegend die extensive Methode.

## 2.1. Zwei Methoden der Analyse des Problems[2]

Zur exemplarischen Darstellung der beiden Methoden soll folgendes einfaches Entscheidungsmodell verwendet werden:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | $k_{11}$ | $k_{12}$ |
| $a_2$ | $k_{21}$ | $k_{22}$ |
| | $p'(s_1)$ | $p'(s_2)$ |

Tab. 4: Konsequenzentabelle eines 2-Aktionen-Problems

Die Informationsaktivität bestehe aus dem Ziehen einer Stichprobe mit Zurücklegen mit dem Umfange n und sei mit $e_n$ symbolisiert.

[1] Loeffel, S. 20 ff.

[2] Vgl. Raiffa, S. 129; Raiffa/Schlaifer, S. 129 ff.

## 2.11. Die Normalform der Analyse

Die Normalform der Analyse beginnt mit einer Betrachtung aller möglichen Entscheidungsregeln. Weil Entscheidungsregeln Angaben darüber sind, welche Aktionen aufgrund von bestimmten Stichprobenresultaten zu wählen sind, hat man sich sinnvollerweise zuerst mit den möglichen Stichprobenausgängen zu beschäftigen. Es sei angenommen, die Stichprobe $e_n$ beziehe sich auf ein qualitatives Merkmal mit zwei Ausprägungen, z.B. x = "Eigenschaft vorhanden" und $\bar{x}$ = "Eigenschaft nicht vorhanden"[1]. In diesem Falle sind $2^n$ Stichprobenresultate möglich, wenn n = 2 die vier Resultate:

1: x, x; 2: x, $\bar{x}$; 3: $\bar{x}$, x; 4: $\bar{x}$, $\bar{x}$.

Die Resultate 2 und 3 unterscheiden sich allerdings nur bezüglich der Reihenfolge des Auftretens der Merkmalsausprägungen. Diese Information ist für den Aktor nicht von Bedeutung. Dadurch reduziert sich die Anzahl der möglichen, den Aktor interessierenden Stichprobenresultate in diesem Falle auf n+1, d.h. R = 0,1,...,n.

Bei der Stichprobe $e_2$ mit ihren drei möglichen Resultaten, hat ein Aktor zwischen den in Tab. 5 aufgeführten Entscheidungsregeln $d_k$ zu wählen:

| R | $d_1$ | $d_2$ | $d_3$ | $d_4$ | $d_5$ | $d_6$ | $d_7$ | $d_8$ |
|---|---|---|---|---|---|---|---|---|
| 0 | $a_1$ | $a_1$ | $a_1$ | $a_2$ | $a_2$ | $a_2$ | $a_1$ | $a_2$ |
| 1 | $a_1$ | $a_1$ | $a_2$ | $a_1$ | $a_2$ | $a_1$ | $a_2$ | $a_2$ |
| 2 | $a_1$ | $a_2$ | $a_1$ | $a_1$ | $a_1$ | $a_2$ | $a_2$ | $a_2$ |

Tab. 5: Mögliche Entscheidungsregel bei einem 2-Aktionen-Problem und einer Stichprobe $e_2$

[1] Aufgrund dieser Voraussetzungen würde sich das Binomial-Modell als stochastisches Modell eignen.

Theoretisch sind bei einer Stichprobe $e_n$, unter den hier getroffenen Voraussetzungen, bei einem 2-Aktionen-Problem $2^{(n+1)}$ Entscheidungsregeln möglich. Wie bei der Anzahl der theoretisch möglichen Stichprobenresultate stellt sich auch hier die Frage, ob der Einbezug all dieser theoretisch möglichen Entscheidungsregeln in das Entscheidungskalkül sinnvoll ist. Diese Frage wird hier verneint. Zur Begründung der Ablehnung soll folgendes Beispiel herangezogen werden:

Es gehe darum, zu entscheiden, ob ein Warenposten angenommen ($a_1$) oder abgelehnt ($a_2$) werden soll, wobei der Entscheid vom Ausschussanteil abhange. Die Stichprobe bezieht sich also auf das qualitative Merkmal "Beschaffenheit eines Werkstückes" mit den beiden Ausprägungen "defekt" bzw. "nicht defekt". Wenn $n=2$ sind drei Resultate möglich, je nachdem ob keines, eines oder beide der zufällig gezogenen Stücke defekt sind. Die Entscheidungsregel $d_3$ der Tab. 5 würde auf das Beispiel bezogen lauten: Wenn kein Stück oder aber beide Stücke defekt sind, soll der Warenposten angenommen werden, nur wenn gerade ein Stück defekt ist, soll er abgelehnt werden. Die Irrationalität dieser Regel ist augenfällig, ebenfalls diejenige der Regel $d_6$, die besagt, dass wenn kein Stück defekt ist, der Posten abgelehnt, wenn aber ein oder gar zwei Stücke defekt sind, angenommen werden soll. Auf unser Beispiel bezogen, kann man diejenigen Entscheidungsregeln als "sinnvoll" definieren, bei denen kein $a_2$ in der Spalte oberhalb eines $a_1$ steht. Die Anzahl der Entscheidungsregeln reduziert sich auf diese Weise auf die in Tab. 6 wiedergegebenen.

| R | $d_1$ | $d_2$ | $d_3$ | $d_4$ |
|---|---|---|---|---|
| 0 | $a_1$ | $a_1$ | $a_1$ | $a_2$ |
| 1 | $a_1$ | $a_1$ | $a_2$ | $a_2$ |
| 2 | $a_1$ | $a_2$ | $a_2$ | $a_2$ |

Tab. 6: Sinnvolle Entscheidungsregel bei einem 2-Aktionen-Problem und einer Stichprobe $e_2$

Mitunter ist es auch üblich, die Entscheidungsregeln mit Hilfe des Begriffs der Annahmezahl k, k=0,1,...,n-1, zu umschreiben:

$d_k$: wenn $R \leq k \rightarrow a_1$

$R > k \rightarrow a_2$

Nach dieser Definition bleiben in unserem Beispiel noch zwei Entscheidungsregeln übrig; $d_1$ und $d_4$ fallen weg. Loeffel schreibt in seiner Broschüre, dass die Auswahl dieser Teilmenge aus der Menge aller möglichen Entscheidungsregeln von der praktischen Seite her intuitiv klar sei, sich aber auch mit dem Fundamentallemma von Neyman-Pearson begründen lasse[1].

Den folgenden Ausführungen wird ebenfalls diese Definition des Begriffs der Entscheidungsregel zugrunde gelegt. (Dabei gilt es zu beachten, dass sich die Indices der Entscheidungsregeln gegenüber Tab. 6 ändern: $d_3$: neu $d_0$; $d_2$: neu $d_1$).

Jede der als zulässig betrachteten Entscheidungsregeln wird nun unter der Annahme, dass ein bestimmter Umweltzustand vorliege, betrachtet. Der folgende Entscheidungsbaum ver-

[1] Loeffel, S. 13

anschauliche dieses Unterfangen am Beispiel der Entscheidungsregel $d_1$:

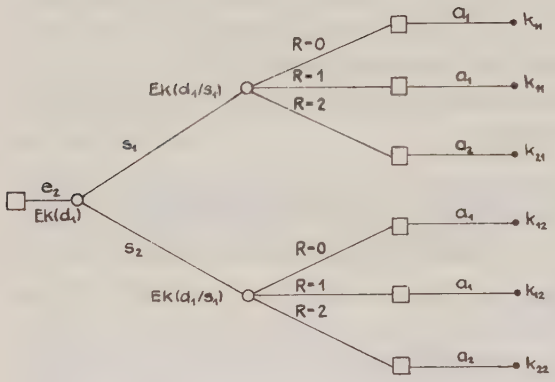


Abb. 11: Entscheidungsbaum für die Entscheidungsregel $d_1$ und n=2 bei Normalform der Analyse

Für die Ermittlung der bedingten Konsequenzenerwartungswerte einer Entscheidungsregel $Ek(d_k/s_j)$ werden die bedingten Wahrscheinlichkeiten $p(R=r/s_j)$ - die Likelihoods - benötigt. Für deren Berechnung ist bekanntlich ein stochastisches Modell - in unserem Falle das Binomial-Modell - erforderlich. Nach der Beziehung (2.1) können die Likelihoods ohne weiteres ermittelt werden.

Die bedingten Konsequenzenerwartungswerte für die verschiedenen Entscheidungsregeln lassen sich zur folgenden

Tabelle zusammenfassen:

| | $Ek\ (d_k/s_1)$ | $Ek\ (d_k/s_2)$ |
|---|---|---|
| $d_0$ | $p(R=0/s_1)k_{11}+p(R=1/s_1)k_{21}+$ $p(R=2/s_1)k_{21}$ | $p(R=0/s_2)k_{12}+p(R=1/s_2)k_{22}+$ $p(R=2/s_2)k_{22}$ |
| $d_1$ | $p(R=0/s_1)k_{11}+p(R=0/s_1)k_{11}+$ $p(R=2/s_1)k_{21}$ | $p(R=0/s_2)k_{12}+p(R=1/s_2)k_{12}+$ $p(R=2/s_2)k_{22}$ |

Tab. 7: Entscheidungsregeln und ihre bedingten Konsequenzenerwartungswerte

Der Konsequenzenerwartungswert einer Entscheidungsregel $d_k$, $Ek(d_k)$ ist gleich der Summe der mit den Apriori-Wahrscheinlichkeiten gewichteten bedingten Konsequenzenerwartungswerte $Ek(d_k/s_j)$:

$$Ek(d_k) = \sum_{j=1}^{J} p'(s_j)Ek(d_k/s_j). \tag{3.16}$$

Werden an Stelle der Konsequenzen Opportunitätskosten verwendet, so ergibt sich kein prinzipieller Unterschied zum geschilderten Verfahren; es gilt sinngemäss:

$$Eo(d_k) = \sum_{j=1}^{J} p'(s_j)Eo(d_k/s_j). \tag{3.17}$$

Optimal ist diejenige Entscheidungsregel, für die gilt

$$Ek(d^*) = \max_k Ek(d_k) \tag{3.18}$$

bzw.
$$Eo(d^*) = \min_k Eo(d_k). \tag{3.19}$$

## 2.12 Die extensive Form der Analyse[1]

Die extensive Methode verzichtet auf eine Ermittlung und Auflistung aller möglichen Entscheidungsregeln und bestimmt die optimale Entscheidungsregel direkt. Der dabei verwendete Entscheidungsbaum - bezogen auf das bekannte Beispiel - hat folgende Struktur:

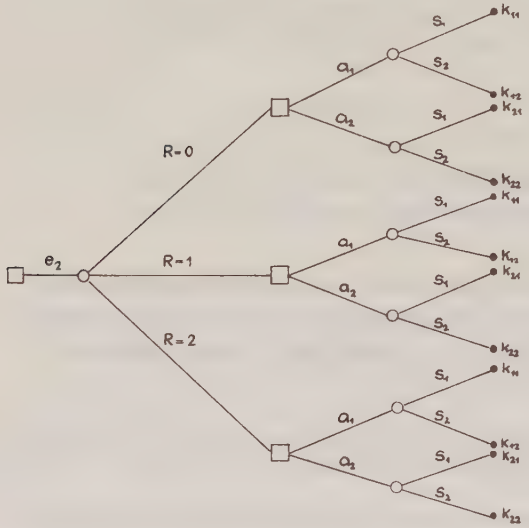


Abb. 12: Entscheidungsbaum für die extensive Form der Analyse

Die optimale Entscheidungsregel wird gefunden, indem man sich von den äussersten Enden des Entscheidbaumes zur Wurzel zurückarbeitet. Dieses Vorgehen heisst "Rückwärtsanalyse" (backwards induction). Es verlangt, dass jedem Zufallsknoten, der einem echten Entscheidungsknoten[2]

[1] Das Kapitel "Einführung in das Bayes'sche Modell des Entscheidens" ist, wie hier ersichtlich wird, auf dieser extensiven Form der Analyse aufgebaut.

[2] Als "echt" wird ein Entscheidungsknoten dann bezeichnet, wenn von diesem Knoten mehrere Aeste ausgehen. Bei der Wurzel des Entscheiungsbaumes in Abb. 11 handelt es sich also nicht um einen echten Entscheidungsknoten; dem ihm nachfolgenden Zufallsknoten, von dem die R-Aeste ausgehen, braucht demzufolge auch kein Erwartungswert zugeschrieben zu werden.

nachgelagert ist, ein Erwartungswert zugeschrieben wird und bei den Entscheidungsknoten die Aktion mit dem günstigeren Erwartungswert gewählt wird. Damit aber diese Erwartungswerte berechnet werden können, müssen die Wahrscheinlichkeiten der Zufallsäste bekannt sein. Es handelt sich dabei, wie aus Abb. 12 ersichtlich ist, um die bedingten Wahrscheinlichkeiten $p''(s_j/R=r)$ - die Aposteriori-Wahrscheinlichkeiten. Diese Aposteriori-Wahrscheinlichkeiten können bekanntlich - wenn die Apriori-Wahrscheinlichkeiten und die Likelihoods bekannt sind - mit Hilfe des Bayes-Theorems ermittelt werden. In unserem Beispiel bietet die Ermittlung der Likelihoods keine Schwierigkeiten, denn es kann natürlich das gleiche stochastische Modell verwendet werden wie bei der Normalform der Analyse.

Wenn ein Aktor all die beschriebenen Operationen durchgeführt hat, kennt er die optimale Entscheidungsregel; er weiss bei jedem möglichen Stichprobenresultat, welche Aktion er zu wählen hat.

Auch bei diesem Verfahren ergeben sich durch die Verwendung von Opportunitätskosten keine Unterschiede.

## 2.13 Vergleich der beiden Methoden

Beide Methoden führen - auf unterschiedlichem Wege, aber vom gleichen Datenmaterial ausgehend - zur selben optimalen Entscheidungsregel. Der Beweis für die Aequivalenz der beiden Methoden findet sich bei Raiffa/Schlaifer[1] oder bei Loeffel[2]. Man kann also ohne Einschränkung jeweils diejenige Methode wählen, die in einer konkreten Entscheidungssituation eine Lösung mit geringerem Aufwand verspricht.

[1] Raiffa/Schlaifer, S. 15

[2] Loeffel, S. 24 f.

Wenn man die beiden Methoden miteinander vergleicht, so fällt auf, dass bei der Normalform der Analyse die für die Bayes'sche Entscheidungstheorie so zentrale formale Relation - das Bayes-Theorem - sowie der Begriff der Aposteriori-Wahrscheinlichkeit gar nicht verwendet werden. Dies gilt nun allerdings nur, wenn die im Abschnitt "Einführung in das Bayes'sche Modell des Entscheidens" erläuterte Normalkonstellation der Wahrscheinlichkeit[1] vorliegt.

Ein weiteres, wesentliches Unterscheidungsmerkmal ist der Zeitpunkt des Einbezuges der Apriori-Wahrscheinlichkeiten in das Kalkül. Während diese Wahrscheinlichkeiten bei der extensiven Methode von Anfang an verwendet werden müssen - sonst könnten die bei dieser Methode zentralen Aposteriori-Wahrscheinlichkeiten nicht ermittelt werden - gehen sie bei der Normalform erst am Schluss ins Kalkül ein. Dies ist ein Umstand, der Sensitivitätsanalysen in bezug auf die Apriori-Wahrscheinlichkeiten erleichtert. Auch wenn die Beurteilung der Apriori-Wahrscheinlichkeiten kontrovers ist - beispielsweise bei Gruppenentscheiden - dürfte dieses Charakteristikum die Normalform bevorteilen[2].

Bei einer Aposteriori-Entscheidung, bei der ein bestimmtes Experiment bereits durchgeführt wurde und ein Resultat vorliegt, hat die extensive Form den Vorteil, dass die optimale Aktion direkt bestimmt werden kann und nicht von sämtlichen möglichen Entscheidungsregeln, wie bei der Normalform, ausgegangen werden muss[3].

---

[1] Vgl. S. 18

[2] Vgl. Raiffa, S. 132 f.; Raiffa/Schlaifer, S. 16

[3] Vgl. Raiffa/Schlaifer, S. 15

## 2.2 Die Entscheidung bei diskreter Apriori-Verteilung

### 2.21. Binomial-Modell

Einer exemplarischen Behandlung soll das in den vorgehenden Kapiteln verwendete Qualitätskontroll-Beispiel dienen; es wird angenommen, dass die Voraussetzungen für eine Anwendung des Binomial-Modells gegeben seien.

Es liege die folgende Konsequenzentabelle vor:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | -10'000 | -60'000 |
| $a_2$ | -40'000 | -40'000 |
| $p'(s_j)$ | 0.4 | 0.6 |

Tab. 8: Konsequenzentabelle für das Qualitätskontroll-Beispiel

$s_1$ symbolisiere einen Ausschussanteil $\Theta_1 = 10\%$; $s_2$ einen Anteil $\Theta_2 = 20\%$; der Stichprobenumfang sein $n = 7$.

Zur Vereinfachung des Problems wird die Konsequenzentabelle in die äquivalente Opportunitätskostentabelle transformiert:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 2 |
| $a_2$ | 3 | 0 |
| | 0.4 | 0.6 |

Tab. 9: Opportunitätskostentabelle für das Qualitätskontroll-Beispiel

Die Anwendung der Normalform der Analyse hat zur Folge, dass die bedingten Opportunitätskosten-Erwartungswerte für n=7 Entscheidungsregeln ermittelt werden müssen. Zu diesem Zweck soll die im Abschnitt über die Normalform anhand eines einfachen Beispiels erläuterte Ermittlung der Opportunitätskosten - Erwartungswerte verallgemeinert werden[1]:

Für eine Entscheidungsregel $d_k$ gilt unter den hier getroffenen Voraussetzungen ($o_{11}$ und $o_{22} = 0$; zwei Umweltzustände)

$$Eo(d_k/s_1) = o_{21} \sum_{r=k+1}^{n} p(R=r/\theta_1) \tag{3.20}$$

$$\text{bzw. } Eo(d_k/s_2) = o_{12} \sum_{r=0}^{k} p(R=r/\theta_2). \tag{3.21}$$

Nach (2.1) ist

$$p(R=r/\theta) = \binom{n}{r} \theta^r (1-\theta)^{n-r}.$$

Eingesetzt in (3.20) bzw. (3.21) erhält man

$$Eo(d_k/s_1) = o_{21} \sum_{r=k+1}^{n} \binom{n}{r} \theta_1^r (1 - \theta_1)^{n-r}$$

und

$$Eo(d_k/s_2) = o_{12} \sum_{r=0}^{k} \binom{n}{r} \theta_2^n (1 - \theta_2)^{n-r}.$$

Mit den Zahlenwerten des Qualitätskontroll-Beispieles ergibt sich exemplarisch für $d_1$:

[1] Vgl. S. 77 und Abb. 11

$$Eo(d_1/s_1) = 3 \sum_{r=2}^{7} \binom{7}{r} 0.1^r \, 0.9^{7-r}$$

$$= 3 \left[1 - \sum_{r=0}^{1} \binom{7}{r} 0.1^r \, 0.9^{7-r}\right]$$

$$= 3 \left[1 - \left(\binom{7}{0} 0.1^0 \, 0.9^7 + \binom{7}{1} 0.1^1 \, 0.9^6\right)\right]$$

$$= 3 \left[1 - (0.4783 + 0.3720)\right]$$

$$Eo(d_1/s_1) = 0.4491$$

bzw. $$Eo(d_1/s_2) = 2 \sum_{r=0}^{1} \binom{7}{r} 0.2^r \, 0.8^{7-r}$$

$$= 2 \left[\binom{7}{0} 0.2^0 \, 0.8^7 + \binom{7}{1} 0.2^1 \, 0.8^6\right]$$

$$= 2 \, (0.2097 + 0.3670)$$

$$Eo(d_1/s_2) = 1.1534 \; .^1$$

Analog wurden die bedingten Opportunitätskosten-Erwartungswerte für die restlichen Entscheidungsregeln gerechnet und schliesslich, nach (3.17), mit den Apriori-Wahrscheinlichkeiten gewichtet (vgl. Tab. 10).

[1] Mit Vorteil werden in diesen Fällen Tabellen mit der Verteilungsfunktion diskret verteilter Zufallsvariablen verwendet, wie sie z.B. bei Peters/Summers, S. 429 ff. zu finden sind.

| $d_k$ | $Eo(d_k/s_1)$ | $Eo(d_k/s_2)$ | $Eo(d_k)$ |
|---|---|---|---|
| $d_0$ | 1.5651 | 0.4194 | 0.8777 |
| $d_1$ | 0.4491 | 1.1534 | 0.8717 |
| $d_2$ | 0.0774 | 1.7040 | 1.0534 |
| $d_3$ | 0.0084 | 1.9334 | 1.1634 |
| $d_4$ | 0.0006 | 1.9908 | 1.1947 |
| $d_5$ | 0 | 1.9994 | 1.1996 |
| $d_6$ | 0 | 2 | 2 |

Tab. 10: Entscheidungsregeln und ihre erwarteten Opportunitätskosten für das Qualitätskontroll-Beispiel

Nach (3.19) ist $d_1$ die optimale Entscheidungsregel. Sie lautet:

wenn $R \leq 1 \rightarrow a_1$;

$R > 1 \rightarrow a_2$.

Eine Sensitivitätsanalyse bezüglich der Apriori-Wahrscheinlichkeit $p'(s_1)$ kann man durchführen, indem man die Apriori-Wahrscheinlichkeiten, mit denen die bedingten Erwartungswerte $Eo(d_k/s_j)$ gewichtet werden, variiert und nach (3.19) die jeweils entsprechende optimale Entscheidungsregel bestimmt. Man erhält dabei in unserem Falle das folgende Resultat:

wenn $0 \leq p'(s_1) \leq 0.39 \rightarrow d_0$

$0.40 \leq p'(s_1) \leq 0.59 \rightarrow d_1$

$0.60 \leq p'(s_1) \leq 0.76 \rightarrow d_2$

$0.77 \leq p'(s_1) \leq 0.88 \rightarrow d_3$

$0.89 \leq p'(s_1) \leq 0.93 \rightarrow d_4$

$0.94 \leq p'(s_1) \leq 1 \rightarrow d_5$

Die Herleitung der optimalen Entscheidungsregel im Falle von mehr als zwei Umweltzuständen bringt keine neuen Probleme mit sich. Bei der Sensitivitätsanalyse entstehen aber Schwierigkeiten, weil nicht mehr wie oben bloss ein Parameter frei festgesetzt werden kann. Eine Sensitivitätsanalyse lässt sich in solchen Fällen auf ceteris-paribus-Basis durchführen. Dabei wird - wie bei unserem Beispiel - ein Parameter variiert, während die anderen konstant gehalten werden[1].

Auf ein gleiches Problem angewandt, geht die extensive Methode von dem in Abb. 13 dargestellten Entscheidungsbaum aus. Bei jedem der mit x bezeichneten Entscheidungsknoten wird die Wahl der optimalen Aktion aufgrund der folgenden Bedingungen getroffen:

wenn (1) $p''(\Theta_2/R=r)\ o_{12} < p''(\Theta_1/R=r)\ o_{21} \rightarrow a_1$

(2) $p''(\Theta_2/R=r)\ o_{12} = p''(\Theta_1/R=r)\ o_{21} \rightarrow$ entweder $a_1$ oder $a_2$

(3) $p''(\Theta_2/R=r)\ o_{12} > p''(\Theta_1/R=r)\ o_{21} \rightarrow a_2$ .

[1] Vgl. Rappaport, S. 178

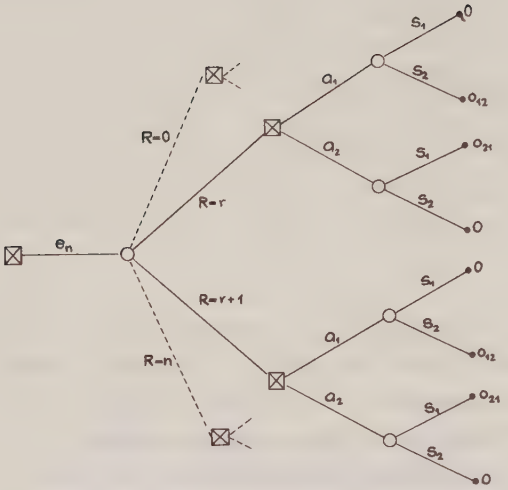


Abb. 13: Entscheidungsbaum für das Qualitätskontroll-Beispiel; extensive Form der Analyse

Die Beziehung (2) erlaubt, den Wert von r zu ermitteln, bei dem die Erwartungswerte der beiden Alternativen gerade gleich gross sind; der grösste ganzzahlige Wert, der kleiner als dieses r ist, ist gleich der Annahmezahl k* der optimalen Entscheidungsregel; dort wird noch $a_1$ gewählt, bei jedem grösseren noch möglichen Resultat aber $a_2$.

Beziehung (2) kann wie folgt umgeformt werden:

$$(4) \quad \frac{p''(\Theta_2/R=r)}{p''(\Theta_1/R=r)} = \frac{o_{21}}{12}$$

Nach dem Bayes-Theorem (1.2) ist aber

$$p''(\theta_2/R{=}r) = \frac{p'(s_2)p(R{=}r/\theta_2)}{p'(s_2)p(R{=}r/\theta_2) + p'(s_1)p(R{=}r/\theta_1)}$$

$$\text{bzw. } p''(\theta_1/R{=}r) = \frac{p'(s_1)p(R{=}r/\theta_1)}{p'(s_1)p(R{=}r/\theta_1) + p'(s_2)p(R{=}r/\theta_2)}$$

Eingesetzt in (4) ergibt sich

$$(5) \quad \frac{p'(s_2)p(R{=}r/\theta_2)}{p'(s_1)p(R{=}r/\theta_1)} = \frac{o_{21}}{o_{12}} .$$

Für die bedingten Wahrscheinlichkeiten $p(R{=}r/\theta_j)$ gilt aufgrund des Binomial-Modells:

$$p(R{=}r/\theta_j) = \binom{n}{r} \theta_j^r (1 - \theta_j)^{n-r}.$$

Eingesetzt in (5):

$$\frac{o_{21}}{o_{12}} = \frac{p'(s_2)\theta_2^r (1-\theta_2)^{n-r}}{p'(s_1)\theta_1^r (1-\theta_1)^{n-r}}.$$

Durch Logarithmieren und nach einigen Umformungen erhält man schliesslich

$$r = \frac{\log[\frac{o_{21}}{o_{12}} \cdot \frac{p'(s_1)}{p'(s_2)}] + n \log [\frac{(1-\theta_1)}{(1-\theta_2)}]}{\log[\frac{\theta_2}{\theta_1} \cdot \frac{(1-\theta_1)}{(1-\theta_2)}]} \cdot \quad (3.22)$$

[1] Für den gleichen Entscheidungstyp und das gleiche stochastische Modell hat Loeffel ebenfalls eine Formel, die eine direkte Herleitung der Annahmezahl der optimalen Entscheidungsregel ermöglicht, hergeleitet. Er geht dabei von einem etwas anderen Ansatz aus und kommt auch zu einem anderen Ergebnis, weil er (u.E. fälschlicherweise) $\theta_2 = 1-\theta_1$ bzw. $\theta_1 = 1-\theta_2$ setzt; Loeffel, S. 29 ff.

Mit den Zahlenwerten unseres Beispiels ist

$$r = \frac{\log\left(\frac{3}{2}\,\frac{0.4}{0.6}\right) + 7\log\left(\frac{0.9}{0.8}\right)}{\log\left(\frac{0.2}{0.1}\,\frac{0.9}{0.8}\right)} = 1.017$$

und demzufolge die Annahmezahl der optimalen Entscheidungsregel $k^*=1$. Ohne Schwierigkeiten kann mit Hilfe der Beziehung (3.22) ebenfalls eine Sensitivitätsanalyse durchgeführt werden.

Der zuerst geschilderte, auf der Normalform der Analyse aufbauende Weg zur Bestimmung der optimalen Entscheidungsregel beruht auf einer systematischen Auswertung sämtlicher Entscheidungsregeln, während beim eben geschilderten Verfahren die Annahmezahl der optimalen Entscheidungsregel direkt bestimmt wird. Das erste Verfahren soll "enumeratives Verfahren", das zweite "analytisches Verfahren" heissen.

## 2.22 Hypergeometrisches Modell

Im letzten Abschnitt wurde angenommen, dass die Voraussetzungen, die eine Anwendung des Binomial-Modells auf das Qualitätskontroll-Beispiel als sinnvoll erscheinen lassen, gegeben seien. Das heisst u.a., dass entweder die Stichprobe mit Zurücklegen gezogen wird oder aber die Grundgesamtheit N, aus der die Stichprobe gezogen wird, gross und der Stichprobenumfang im Verhältnis zur Grundgesamtheit relativ klein ist ($n/N<0.1$; $N\geq 60$)[1]. Sind diese Voraussetzungen nicht gegeben, so muss das hypergeometrische Modell angewendet werden. Im folgenden wird

[1] Sachs, S. 175

das Qualitätskontroll-Beispiel unverändert übernommen, die Stichprobe werde aber ohne Zurücklegen gezogen und die Grundgesamtheit sei N=40.

Im Vergleich zu dem unter dem Titel des Binomial-Modells geschilderten Verfahren ergibt sich kein prinzipieller Unterschied. In die Gleichungen (3.20) bzw. (3.21) wird nach (2.2)

$$p(R=r/\Theta) = \frac{\binom{\Theta N}{r}\binom{(1-\Theta)N}{n-r}}{\binom{N}{n}}$$

eingesetzt. Dadurch erhält man

$$Eo(d_k/s_1) = o_{21} \sum_{r=k+1}^{n} \frac{\binom{\Theta_1 N}{r}\binom{(1-\Theta_1)N}{n-r}}{\binom{N}{n}}$$

und

$$Eo(d_k/s_2) = o_{12} \sum_{r=0}^{k} \frac{\binom{\Theta_2 N}{r}\binom{(1-\Theta_2)N}{n-r}}{\binom{N}{n}}$$

Für $d_1$ ergibt sich beispielsweise:

$$Eo(d_1/s_1) = 3 \sum_{r=2}^{7} \frac{\binom{0.1\cdot 40}{r}\binom{0.9\cdot 40}{7-r}}{\binom{40}{7}}$$

$$= 3\left[1 - \sum_{r=0}^{1} \frac{\binom{0.1\cdot 40}{r}\binom{0.9\cdot 40}{7-r}}{\binom{40}{7}}\right]$$

$$= 3\left[1-\left(\frac{\binom{4}{0}\binom{36}{7}}{\binom{40}{7}} + \frac{\binom{4}{1}\binom{36}{6}}{\binom{40}{7}}\right)\right]$$

$$= 3\,[1-(0.4478 + 0.4179)]$$

$$Eo(d_1/s_1) = 0.4029.$$

Analog erhält man

$$Eo(d_1/s_2) = 1.1386.$$

Wie Tab. 11 zeigt, ist auch unter den hier getroffenen Annahmen $d_1$ die optimale Entscheidungsregel. Auf eine Sensitivitätsanalyse soll verzichtet werden, es würden sich dabei keine neuen Probleme stellen.

| $d_k$ | $Eo(d_k/s_1)$ | $Eo(d_k/s_2)$ | $Eo(d_k)$ |
|---|---|---|---|
| $d_0$ | 1.6566 | 0.3610 | 0.8792 |
| $d_1$ | 0.4029 | 1.1386 | 0.8443 |
| $d_2$ | 0.0390 | 1.7434 | 1.0616 |
| $d_3$ | 0.0012 | 1.9594 | 1.1761 |
| $d_4$ | 0 | 1.9970 | 1.1982 |
| $d_5$ | 0 | 2 | 1.2000 |
| $d_6$ | 0 | 2 | 1.2000 |

Tab. 11: Entscheidungsregeln und ihre Opportunitätskosten-Erwartungswerte für das Qualitätskontroll-Beispiel mit einer Stichprobe ohne Zurücklegen

Mit der gleichen Vorgehensweise wie im letzten Abschnitt kann auch für das hypergeometrische Modell eine Formel, die eine direkte Bestimmung von k* auf dem analytischen Wege ermöglicht, hergeleitet werden. Der einzige Unterschied zum geschilderten Verfahren besteht wiederum darin, dass die bedingten Wahrscheinlichkeiten $p(R=r/\Theta)$ durch das hypergeometrische Modell bestimmt werden.

Man gelangt auf diesem Wege zur folgenden, etwas kompliziert anmutenden Gleichung:

$$\frac{o_{21}\, p'(s_1)\; a!\;\; (N-a)!}{o_{12}\, p'(s_2)\; b!\;\; (N-b)!} = \frac{(a-r)!\;\; (N-a-n+r)!}{(b-r)!\;\; (N-b-n+r)!} \qquad (3.23)$$

wobei $a=\theta_1 N$ und $b=\theta_2 N$.

Trotz ihrer etwas abschreckenden Form und dem Umstand, dass sich die Gleichung nicht nach r auflösen lässt, ermöglicht diese Beziehung eine mühelose Bestimmung der Annahmezahl der optimalen Entscheidungsregel. Es sei dies mit den Werten unseres Beispiels demonstriert.

Dabei ist

$$\frac{o_{21}\, p'(s_1)\; a!\;\; (N-a)!}{o_{12}\, p'(s_2)\; b!\;\; (n-b)!} = 841.5.$$

Ferner sei

$$\frac{(a-r)!\;\; (N-a-n+r)!}{(b-r)!\;\; (N-b-n+r)!} = A.$$

Wenn $r = 0 \rightarrow A = 339.3$

$r = 1 \rightarrow A = 783$

$r = 2 \rightarrow A = 2'097.67,$

d.h., $1<r<2$ und demzufolge ist $k^*=1$.

(Es sei hier nochmals explizit darauf hingewiesen, dass sowohl die Beziehung (3.23) als auch (3.22) nur für Entscheidungssituationen mit bloss zwei Umweltzuständen gelten.)

Schliesslich noch eine abschliessende Bemerkung zum hypergeometrischen Modell: Die Anwendung dieses Modells bietet weitaus grössere rechentechnische Schwierigkeiten als das Binomial-Modell. Dazu kommt, dass die hypergeometrische Verteilung in der gängigen statistischen Literatur nicht - und dies wieder im Gegensatz zur Binomialverteilung - tabelliert ist.

Im Hinblick auf diese Schwierigkeiten wird man argumentieren, dass die Bedingungen, die eine Anwendung des hypergeometrischen Modells angezeigt erscheinen lassen, ja jederzeit so modifiziert werden könnten, dass das Binomial-Modell praktikabel ist; konkret: dass einfach die Stichprobe mit Zurücklegen gezogen werden müsse. Dies ist aber nicht in jedem Falle möglich, dann nämlich nicht, wenn ein gezogenes Stück bei der Stichprobe zerstört wird, wie dies in Qualitätskontrollbeispielen bisweilen der Fall ist.

## 2.23 Poisson-Modell

Auch an den Anfang dieses Abschnittes soll ein kurzer Fall gestellt werden[1]:

Der Inhaber einer Automobil-Handlung muss einen neuen Autoverkäufer einstellen. Aus einer Reihe von Bewerbern bleibt schliesslich derjenige übrig, der einen kurzen Eignungstest als einziger positiv absolviert hat. Aufgrund dieses Testes schätzt der Inhaber die Chance, dass es sich bei diesem Bewerber um einen guten Autoverkäufer handelt, gleich 0.6 ein. Der Mann soll probeweise für 24 Arbeits-

[1] Das Beispiel wurde - leicht modifiziert - von Winkler übernommen, der allerdings eine andere Lösungsmethodik anwendet; Winkler, S. 90 ff.

tage eingestellt werden. Aufgrund seiner Leistung - d.h. der Anzahl verkaufter Autos - soll dann der definitive Entscheid über die Einstellung fallen.

Es lässt sich folgende Opportunitätskostentabelle zusammenstellen:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 2 |
| $a_2$ | 3 | 0 |
| $p(s_j)$ | 0.6 | 0.4 |

Tab. 12: Opportunitätskostentabelle für das Autoverkäufer-Beispiel

Dabei sind

$a_1$ : Bewerber definitiv einstellen;

$a_2$ : Bewerber nicht einstellen;

$s_1$ : Der Bewerber ist ein guter Autoverkäufer, d.h. er verkauft pro Woche im Schnitt 2 Autos oder pro Tag $\lambda_1 = 0.4$ Autos;

$s_2$ : Der Bewerber ist ein schlechter Autoverkäufer, d.h. er verkauft pro Woche im Schnitt nur ein Auto oder $\lambda_2 = 0.2$ Autos pro Tag.

Die Stichprobe bezieht sich in diesem Falle auf ein Zeitintervall t=24, während dieses Intervalls sind theoretisch unbegrenzt viele Erfolge - d.h. Autoverkäufe des probeweise eingestellten Verkäufers - möglich.

Der Entscheidungsbaum in Abb. 14 - auf der Normalform der Analyse basierend - macht deutlich, dass dieses Entschei-

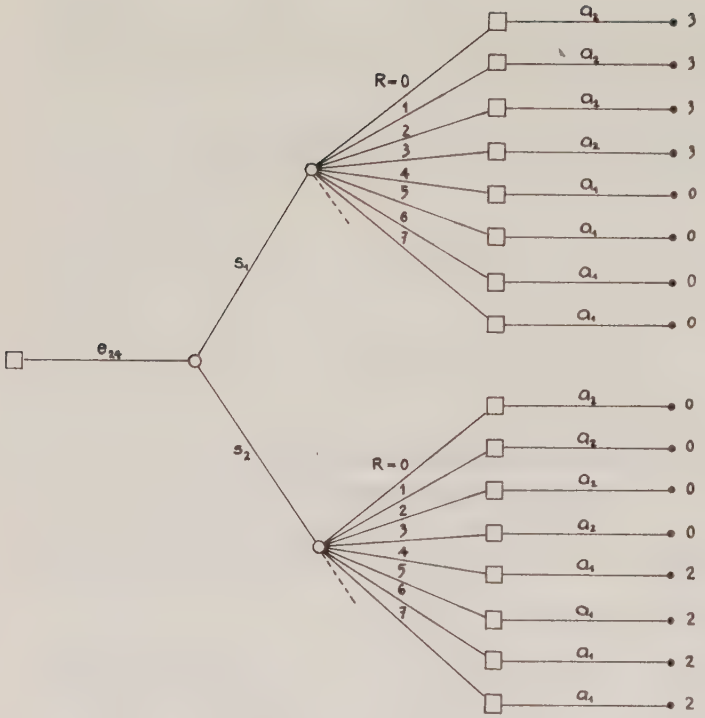


Abb. 14: Entscheidungsbaum für das Autoverkäufer-Beispiel - exemplarisch für $d_4$

dungsproblem eine etwas andere Struktur aufweist als die bisher behandelten Fälle. Nebst dem Umstand, dass die Anzahl der günstigen Ereignisse nach oben nicht begrenzt ist - d.h. auch, dass eine unbegrenzte Anzahl von Entscheidungsregeln möglich ist - hat auch das Verhältnis zwischen Resultat und Annahmezahl einen anderen Charakter; eine Entscheidungsregel wird hier durch die Bezie-

hung

$$d_k : \text{ wenn } R \geq k \rightarrow a_1$$
$$R < k \rightarrow a_2$$

definiert. Die Beziehungen (3.20) und (3.21) müssen deshalb wie folgt geändert werden:

$$Eo(d_k/s_1) = o_{21} \sum_{r=0}^{k-1} p(R=r/\lambda_1) \tag{3.24}$$

und $$Eo(d_k/s_2) = o_{12} \sum_{r=k}^{\infty} p(R=r/\lambda_2) \tag{3.25}$$

Der Umstand, dass die Ereignisse über einem Zeitkontinuum auftreten, lässt die Verwendung des Poisson-Modells für die Ermittlung der Likelihoods angezeigt erscheinen und es gilt nach (2.4)

$$p(r/\lambda t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}.$$

Eingesetzt in (3.24) und (3.25) ergibt sich

$$Eo(d_k/s_1) = o_{21} \sum_{r=0}^{k-1} \frac{(\lambda_1 t)^r}{r!} e^{-\lambda_1 t}$$

und $$Eo(d_k/s_2) = o_{12} \sum_{r=k}^{\infty} \frac{(\lambda_2 t)^r}{r!} e^{-\lambda_2 t}.$$

So ist beispielsweise

$$Eo(d_4/s_1) = 3 \sum_{r=0}^{3} \frac{(0.4 \cdot 24)}{r!} e^{-(0.4 \cdot 24)}$$

$$= 3\,[e^{-9.6} + 9.6e^{-9.6} + \frac{9.6^2}{2}e^{-9.6} + \frac{9.6^3}{6}e^{-9.6}]$$

$$= 3\ (0.0001 + 0.0007 + 0.0031 + 0.0100)$$

$$Eo(d_4/s_1) \doteq 0.0417.$$

$$Eo(d_4/s_1) = 2\,[1 - \sum_{r=0}^{3} \frac{(0.2 \cdot 24)^r}{r!} e^{-(0.2 \cdot 24)}]$$

$$= 1.4116.$$

Tab. 13 fasst die analog ermittelten erwarteten Opportunitätskosten für weitere Entscheidungsregeln zusammen. Optimale Entscheidungsregel d* ist $d_6$.

| $d_k$ | $Eo(d_k/s_1)$ | $Eo(d_k/s_2)$ | $Eo(d_k)$ |
|---|---|---|---|
| $d_1$ | 0.0003 | 1.9832 | 0.7934 |
| $d_2$ | 0.0024 | 1.9042 | 0.7631 |
| $d_3$ | 0.0117 | 1.7146 | 0.6929 |
| $d_4$ | 0.0417 | 1.4116 | 0.5897 |
| $d_5$ | 0.1137 | 1.0472 | 0.4871 |
| $d_6$ | 0.2517 | 0.6978 | 0.4301 |
| $d_7$ | 0.4725 | 0.4182 | 0.4508 |
| $d_8$ | 0.7755 | 0.2264 | 0.5559 |
| $d_9$ | 1.1391 | 0.1114 | 0.7280 |
| . | . | . | . |
| . | . | . | . |

Tab. 13: Entscheidungsregeln und ihre erwarteten Opportunitätskosten für das Autoverkäufer-Beispiel

Die Herleitung von k* auf analytischem Weg erfolgt auf die gleiche Art und Weise wie beim Binomial- bzw. Hypergeometrischen Modell. Wegen der modifizierten Definition der Entscheidungsregel ist aber nicht der grösste ganzzahlige Wert, der kleiner ist als das durch die Formel bestimmte r, sondern der kleinste ganzzahlige Wert, der grösser als dieses r ist, die gesuchte Annahmezahl k*.

Die Formel lautet

$$r = \frac{\ln[\frac{o_{21}p'(s_1)}{o_{12}p'(s_2)}] + t(\lambda_2-\lambda_1)}{\ln(\frac{\lambda_2}{\lambda_1})}$$

Mit den Zahlenwerten unseres Beispiels ist

$$r = \frac{\ln 2.25 + 24(-0.2)}{\ln 0.5} = 5.76$$

und k* = 6.

2.24 Normal-Modell

Eine Nahrungsmittelfirma habe folgendes Problem zu lösen[1]:

Die Werbeabteilung hat für das Spitzenprodukt der Firma - Instant-Kaffee - ein Konzept für eine neue Verpackung entwickelt, die für den Käufer echte Vorteile bezüglich Frischhaltung bieten soll. Die neue Verpackung käme, was den Materialaufwand anbelangt, nicht teuerer zu stehen als die bisherige, hingegen wären für neue Verpackungsanlagen Investititonen von insgesamt Fr. 500'000.-- notwendig. Der mit dem Entscheid über die Einführung der neuen Verpackung betraute Marketingleiter rechnet damit, dass die Konkurrenz innert zweier Jahre auf den Vorstoss seiner Firma reagieren wird und dass sich von diesem Zeit-

[1] Die Problemstellung wurde - stark abgewandelt - von Schlaifer übernommen; Schlaifer, S. 227 f.

punkt an die Marktanteile wieder auf der bisherigen Höhe einpendeln werden. Vom Marktvolumen glaubt man, dass es sich in absehbarer Zeit kaum verändern wird. Der Instant-Kaffee bringt pro kg abgesetzter Menge Fr. 1.-- Deckungsbeitrag ein, d.h., dass die 20'000 Detailhandelsgeschäfte, die den Kaffee verkaufen, während zweier Jahre im Schnitt pro Monat ungefähr 1.05 kg mehr absetzen müssen, damit die Investition rentiert. (Zur Vereinfachung des Problems wird dabei von einer Diskontierung der Einnahmen abgesehen.)

Der monatliche Durchschnittsabsatz aller Geschäfte war bisher normalverteilt mit $\mu = 50$ kg und $\sigma = 15$ kg; saisonale Schwankungen traten keine auf.

Der Marketingleiter stellt sich folgende Konsequenzentabelle auf:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 460'000 | -500'000 |
| $a_2$ | 0 | 0 |
| $p'(s_j)$ | 0.6 | 0.4 |

Tab. 14: Konsequenzentabelle für das Verpackungs-Beispiel

Dabei sind

$a_1$: neue Verpackung einführen;
$a_2$: neue Verpackung nicht einführen;
$s_1$: die Marktlage ist günstig: der durchschnittliche Monatsabsatz steigt um 2 kg : $\mu_1 = 52$ kg;
$s_2$: die Marktlage ist ungünstig: Es kann nicht mehr als bisher verkauft werden: $\mu_2 = 50$ kg.

Bevor der Entscheid gefällt wird, soll in 100 zufällig bestimmten Geschäften während je eines Monats ein Test mit der neuen Verpackung durchgeführt werden.

Wie soll der Marketingleiter aufgrund der Testresultate entscheiden?

Zunächst soll die Konsequenzentabelle in die - ihr äquivalente - Opportunitätskostentabelle transformiert werden:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 25 |
| $a_2$ | 23 | 0 |
| $p'(s_j)$ | 0.6 | 0.4 |

Tab. 15: Opportunitätskostentabelle für das Verpackungsproblem

Eine Entscheidungsregel kann in diesem Falle folgendermassen umschrieben werden: wenn der monatliche Durchschnittsabsatz der Testgeschäfte einen bestimmten kritischen Wert übersteigt, soll die neue Verpackung eingeführt, andernfalls das bisherige System beibehalten werden. Dieser kritische Wert soll wie bisher Annahmezahl k genannt werden, er bezeichnet hier aber nicht mehr eine Anzahl günstiger bzw. ungünstiger Ereignisse, sondern eine bestimmte Höhe des Stichprobenmittelwertes m in kg.

Es gilt für $d_k$: Wenn $M \geq k \rightarrow a_1$

$M < k \rightarrow a_2$.

Die Problemstellung sei mit einem Entscheidungsbaum - nach der Normalform der Analyse - für die Entscheidungsregel $d_k$ charakterisiert:

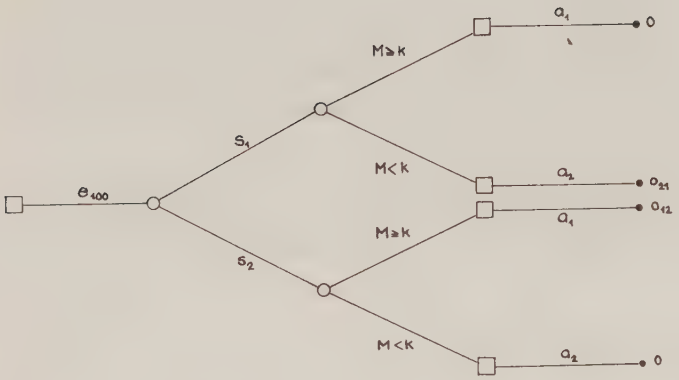


Abb. 15: Entscheidungsbaum für das Verpackungs-Beispiel

Geht man von der Hypothese aus, dass die Marktlage günstig bzw. $s_1$ wahr ist, so bezeichnet die bedingte Wahrscheinlichkeit (Likelihood) $p(M < k/s_1)$ die Wahrscheinlichkeit, mit der die richtige Hypothese - aufgrund des Stichprobenresultates - abgelehnt wird. Diese Wahrscheinlichkeit wird in der Statistik üblicherweise "α-Fehler" ($p_\alpha$) bzw. "Fehler 1. Art" genannt.

Die Wahrscheinlichkeit, dass eine falsche Hypothese angenommen wird, heisst demgegenüber "β-Fehler" ($p_\beta$) oder "Fehler 2. Art". Aufgrund der hier formulierten Ausgangshypothese handelt es sich dabei um $p(M \geq k/s_2)$.

Diese Ausgangshypothese heisst auch "Nullhypothese", und es steht ihr jeweils eine Alternativhypothese - die in unserem Falle "die Marktlage ist ungünstig" lauten würde - gegenüber. Wie man leicht sieht, werden α- und β-Fehler gerade vertauscht, wenn man "die Marktlage ist ungünstig" bzw. "$s_2$ ist wahr" als Nullhypothese setzt: $p(M \geq k/s_2)$ be-

zeichnet in diesem Falle den α-Fehler und $p(M<k/s_1)$ den β-Fehler[1].

Im folgenden sei "$s_1$ ist wahr" als Nullhypothese gesetzt. Der bedingte Opportunitätskostenerwartungswert $d_k$ kann nun unter Verwendung der eingeführten Begriffe wie folgt allgemein umschrieben werden:

$$Eo(d_k/s_1) = o_{21}\, p_\alpha \tag{3.27}$$

bzw. $$Eo(d_k/s_2) = o_{12}\, p_\beta \tag{3.28}$$

und für den Opportunitätskosten-Erwartungswert gilt nach (3.17)

$$Eo(d_k) = p'(s_1)\, o_{21}\, p_\alpha + p'(s_2)\, o_{21}\, p_\beta \tag{3.29}$$

Am Beispiel von $d_{51.2}$ - der kritische Wert k beträgt also 51.2 kg - soll nun gezeigt werden, wie $p_\alpha$ und $p_\beta$ mit Hilfe des Normal-Modells ermittelt werden können.

Nach dem zentralen Grenzwertsatz sind die Stichprobenmittelwerte m normalverteilt mit dem Erwartungswert $\mu_m=\mu$ und der Varianz $\sigma_m^2=\frac{\sigma^2}{n}$ [2]. Wenn angenommen wird, dass sich die

---

[1]Es sei hier noch vermerkt, dass gemäss diesem Sprachgebrauch - je nach der Formulierung der Nullhypothese - die Summen der bedingten Wahrscheinlichkeiten der Beziehungen

(3.20): $\sum_{r=k+1}^{n} p(R=r/\theta_1)$, (3.21): $\sum_{r=0}^{k} p(R=r/\theta_2)$, (3.24): $\sum_{r=0}^{k-1} p(R=r/\lambda_1)$,

(3.25): $\sum_{r=k}^{\infty} p(R=r/\lambda_2)$ ebenfalls α- oder β-Fehler genannt werden können.

[2]Vgl. S. 30

Varianz der Grundgesamtheit durch die Einführung der neuen Verpackung nicht verändert, gilt für das Verpakkungs-Beispiel $\sigma_m^2 = 2.25$ kg bzw. $\sigma_m = 1.5$ kg.

α-Fehler und β-Fehler können grafisch wie folgt dargestellt werden:

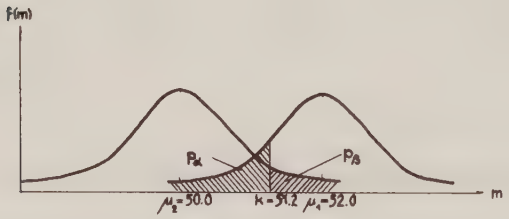


Abb. 16: α- und β-Fehler beim Verpackungs-Beispiel

Mittels der in jedem statistischen Buch zu findenden Tabelle über die standardisierte Normalverteilung können $p_\alpha$ und $p_\beta$ ohne weiteres ermittelt werden. Dem Wert k entspricht bei der standardisierten Normalverteilung der Achsenabschnitt $z_\alpha$ bzw. $z_\beta$ und es gilt nach (2.7)

$$z_\alpha = \frac{k-\mu_1}{\sigma_m} = \frac{51.2 - 52}{1.5} = -0.53$$

$$\text{bzw. } z_\beta = \frac{k-\mu}{\sigma_m} = \frac{51.2 - 52}{1.5} = 0.80.$$

Einer der erwähnten Tabellen lassen sich die zugehörigen Werte für $p_\alpha$ und $p_\beta$ entnehmen:

wenn $z_\alpha = -0.53$, dann ist $p_\alpha = 0.2981$;

wenn $z_\beta = 0.80$, dann ist $p_\beta = 0.2119$.

Nach (3.29) ist

$Eo(d_{51.2}) = 0.6 \cdot 23 \cdot 0.2981 + 0.4 \cdot 25 \cdot 0.2119 = 6.2328.$

Offensichtlich ist auch bei dieser Problemlage die Anzahl möglicher Entscheidungsregeln nicht begrenzt. Indem man versucht, das Intervall, in dem sich die optimale Entscheidungsregel befindet, abzugrenzen und dann schrittweise zu verkleinern, kann man sich aber der optimalen Entscheidungsregel mit vertretbarem Aufwand nähern. Tab. 16 ist nach diesem Prinzip aufgebaut. Mit einer Genauigkeit von einer Stelle nach dem Komma wurde nach sechs Schritten $d_{50.7}$ als optimale Entscheidungsregel ermittelt.

| $d_k$ | $Eo(d_k/s_1)$ | $Eo(d_k/s_2)$ | $Eo(d_k)$ |
|---|---|---|---|
| $d_{51.2}$ | 6.8563 | 5.2975 | 6.2328 |
| $d_{50.5}$ | 3.6501 | 9.1725 | 5.8971 |
| $d_{51.0}$ | 5.7822 | 6.2850 | 5.9833 |
| $d_{50.6}$ | 4.0526 | 8.6150 | 5.8776 |
| $d_{50.8}$ | 4.8737 | 7.4525 | 5.9052 |
| $d_{50.7}$ | 4.4206 | 7.9800 | 5.8444 |

Tab. 16: Entscheidungsregeln und ihre erwarteten Opportunitätskosten für das Verpackungs-Beispiel

Die Bestimmung der Annahmezahl k* auf analytischem Weg geht von der Beziehung (3.19) aus, nach der diejenige Entscheidungsregel optimal ist, für die gilt

(1) $Eo(d^*) = \min_k Eo(d_k)$.

Es gelte ferner $\mu_1 > \mu_2$.

Nach (3.29) ist

(2) $Eo(d_k) = p'(s_1)\, o_{21}\, p_\alpha + p'(s_2)\, o_{12}\, p_\beta$

und nach (2.6) gilt

$$p_\alpha = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_\alpha} e^{-\frac{1}{2}z^2}\, dz$$

bzw. $$p_\beta = \frac{1}{\sqrt{2\pi}} \int_{z\beta}^{\infty} e^{-\frac{1}{2}z^2}\, dz$$

wobei $z_\alpha = \frac{k-\mu_1}{\sigma_m}$

und $z_\beta = \frac{k-\mu_2}{\sigma_m}$ .

Eingesetzt in (2) ergibt sich

$$Eo(d_k) = p'(s_1) o_{21} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{k-\mu_1}{\sigma_m}} e^{-\frac{1}{2}z^2}\, dz + p'(s) o_{12} \frac{1}{\sqrt{2\pi}} \int_{\frac{k-\mu_2}{\sigma_m}}^{\infty} e^{-\frac{1}{2}z^2}\, dz.$$

Die notwendige Bedingung (1) ist erfüllt, wenn

$$(3)\quad \frac{d}{dk}\left[p'(s_1)o_{21}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{k-\mu_1}{\sigma_m}} e^{-\frac{1}{2}z^2}\,dz + p'(s_2)o_{12}\frac{1}{\sqrt{2\pi}}\int_{\frac{k-\mu_2}{\sigma_m}}^{\infty} e^{-\frac{1}{2}z^2}\,dz\right] = 0$$

gilt.

Die Auflösung des Ausdruckes (3) ergibt nach einigen Schritten die Formel für die Annahmezahl k* der optimalen Entscheidungsregel[1]

$$k^* = \frac{\frac{\mu_1^2-\mu_2^2}{\sigma_m^2} - 2\ln\frac{o_{21}p'(s_1)}{o_{12}p'(s_2)}}{\frac{2(\mu_1-\mu_2)}{\sigma_m^2}}\,.$$

Weil aber $\sigma_m^2 = \frac{\sigma^2}{n}$,

$$\text{gilt } k^* = \frac{\frac{n(\mu_1^2-\mu_2^2)}{\sigma^2} - 2\ln\frac{o_{21}p'(s_1)}{o_{12}p'(s_2)}}{\frac{2n(\mu_1-\mu_2)}{\sigma^2}} \qquad (3.30)$$

Weil die zweite Ableitung

$$\frac{d^2Eo(d_k)}{dk^2} = \frac{2(\mu_1-\mu_2)}{\sigma_m^2}$$

[1] Für die vollständige Herleitung s. Anhang, S. 170 ff.

und nach unserer Definition

$$\mu_1 > \mu_2,$$

ist $\dfrac{d^2 Eo(d_k)}{dk^2} > 0,$

also auch die hinreichende Bedingung erfüllt.

Mit den Zahlenwerten unseres Beispiels erhält man

$$k^* = \frac{\frac{52^2-50^2}{2.25} - 2 \ln \frac{23 \cdot 0.6}{25 \cdot 0.4}}{\frac{2(52-50)}{2.25}} = 50.637.$$

Die Diskrepanz zu dem mit Hilfe des enumerativen Verfahrens gewonnenen Resultat ist darauf zurückzuführen, dass beim ersten Verfahren die z-Werte jeweils auf zwei Stellen nach dem Komma gerundet wurden.

Einen etwas anderen Aspekt des Normal-Modells berührt der Fall der PAN ATLANTIC Airways[1]:

Eine Luftfahrgesellschaft, die PAN ATLANTIC Airways, hat zu entscheiden, ob eine neue Fluglinie zu eröffnen sei. Der Entscheid hängt vom zu erwartenden Auslastungsgrad ab. Der Break-even-Punkt liegt bei einer Auslastung von 65%. Zur Stützung des Entscheides soll eine Befragung von 200 potentiellen Kunden durchgeführt werden. Die Befragung ist so konzipiert, dass die Antworten zwei Kategorien - "starkes Interesse an der neuen Fluglinie" und "kein starkes Interesse an der neuen Fluglinie" - zugeordnet werden können. Eine Auslastung von mehr als 65% kann als sicher

[1] Vgl. Gygi/Weibel

angenommen werden, wenn mindestens 70% sämtlicher potentiellen Kunden "starkes Interesse" bekunden.

Die Opportunitätskostentabelle habe folgendes Aussehen:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 5 |
| $a_2$ | 9 | 0 |
| $p'(s_j)$ | 0.6 | 0.4 |

Tab. 17: Opportunitätskostentabelle für das Problem der PAN ATLANTIC Airways

wobei $a_1$ : neue Fluglinie eröffnen;
$a_2$ : neue Fluglinie nicht eröffnen;
$s_1$ : die Marktlage ist günstig, 80% = $\Theta_1$ sämtlicher potentieller Kunden haben starkes Interesse an der neuen Linie;
$s_2$ : die Marktlage ist ungünstig, nur 70% = $\Theta_2$ haben starkes Interesse an der neuen Linie.

Die Nullhypothese, welche mittels der Befragung getestet werden soll, sei: "Die Marktlage ist ungünstig" oder "$s_2$ ist wahr", ihr steht die Alternativhypothese "$s_1$ ist wahr" oder "die Marktlage ist günstig" gegenüber. Wenn der Anteil q jener Befragten, welche "starkes Interesse" bekunden, grösser oder gleich einem bestimmten kritischen Wert k ist, so wird die Nullhypothese verworfen, d.h. $a_1$ gewählt, im umgekehrten Falle aber $a_2$. Weil es sich beim kritischen Wert k jetzt um einen Anteilswert handelt, soll in diesem Zusammenhang von der "Anteilsproportion" die Rede sein. Es gilt also

wenn $Q \geq k \rightarrow a_1$

$Q < k \rightarrow a_2$.

Die Wahrscheinlichkeit, dass die richtige Hypothese (Nullhypothese) verworfen wird, ist

$$p_\alpha = p(Q \geq k/s_2)$$

bzw. die Wahrscheinlichkeit, dass eine falsche Hypothese (die Alternativhypothese) angenommen wird, ist

$$p_\beta = p(Q < k/s_1)$$ (vgl. Abb. 17).

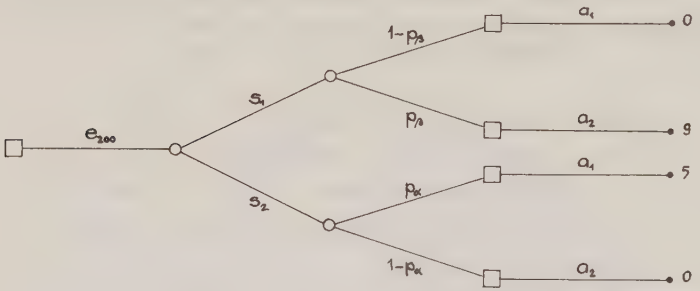


Abb. 17: Entscheidungsbaum für den Fall der PAN ATLANTIC Airways

Für die Ermittlung der Likelihoods $p_\alpha$ bzw. $p_\beta$ kommt in dieser Situation das Binomial- oder Normal-Modell - letzteres als Approximation des Binomial-Modells[1] - in Frage.

Nach dem Normal-Modell, sind die Stichprobenanteilswerte $(\theta, \sqrt{\frac{\theta(1-\theta)}{n}})$ normalverteilt[2].

Es gilt nun zu beachten, dass die Nullhypothese hier anders

[1] Vgl. S. 28

[2] A.a.O.

formuliert wurde, als beim Verpackungs-Beispiel. Im Falle der PAN ATLANTIC Airways ist es die für den Aktor ungünstige Situation, die zur Nullhypothese erklärt wird, beim Verpackungsproblem war es gerade umgekehrt.

In der Statistik gilt die Regel, dass die Nullhypothese aufgestellt wird, um verworfen zu werden. Für unsere Zwecke, die Ermittlung der Annahmeproportion der optimalen Entscheidungsregel, spielt es aber keine Rolle, welche der Situationen zur Nullhypothese erhoben wird; beide Vorgehensweisen führen zum selben Resultat.

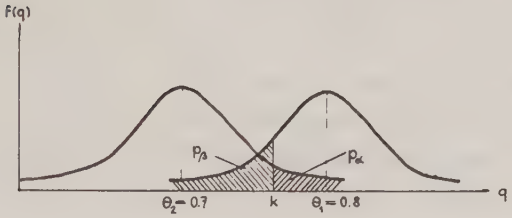


Abb. 18: α- und β-Fehler beim Fall der PAN ATLANTIC Airways

Abb. 18 macht deutlich, dass α- und β-Fehler wegen dieser unterschiedlichen Formulierung der Nullhypothese im Vergleich zum Verpackungsproblem (vgl. Abb. 16) gerade vertauscht sind. Zudem kompliziert sich die Situation hier noch, weil die Standardabweichungen der beiden Normalverteilungen nicht gleich gross sind.

Es sind

$$p_\alpha = \frac{1}{\sqrt{2\pi}} \int_{z_\alpha}^{\infty} e^{-\frac{1}{2}z^2} dz$$

und

$$p_\beta = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_\beta} e^{-\frac{1}{2}z^2} dz,$$

wobei

$$z_\alpha = \frac{k-\theta_2}{\sigma_2} \text{ und } \sigma_2 = \sqrt{\frac{\theta_2(1-\theta_2)}{n}} = 0.03240$$

bzw.

$$z_\beta = \frac{k-\theta_1}{\sigma_1} \text{ und } \sigma_1 = \sqrt{\frac{\theta_1(1-\theta_1)}{n}} = 0.02828 .$$

Im übrigen wird bei der Herleitung der Formel für die direkte Bestimmung der Annahmeproportion der optimalen Entscheidungsregel wie oben vorgegangen[1]. Man gelangt auf diese Weise zur folgenden Beziehung

$$k_{1,2} = \frac{na + \sqrt{n^2a^2 - nb\left[na + 2 \ln \frac{o_{21}p'(s_1)\sqrt{\theta_2(1-\theta_2)}}{o_{12}p'(s_2)\sqrt{\theta_1(1-\theta_1)}}\right]}}{nb}, \qquad (3.31)$$

mit $a = \frac{1}{1-\theta_2} - \frac{1}{1-\theta_1}$ und $b = \frac{1}{\theta_2(1-\theta_2)} - \frac{1}{\theta_1(1-\theta_1)}$ .

Dabei sind zwei Lösungen möglich, weil k aus einer quadratischen Gleichung hergeleitet werden muss. Für k* gilt

$$0 \leq k^* \leq 1.$$

Wenn sowohl $k_1$ als auch $k_2$ diese Bedingung erfüllen, so

[1] Vgl. Anhang, S. 172

ist derjenige Wert zu wählen, der den geringeren Opportunitätskosten-Erwartungswert $Eo(d_k)$ zur Folge hat.

Mit den Werten unseres Beispiels ergibt sich

$$k_1 = 0.7432 \quad \text{und} \quad k_2 = 1.4968;$$

die Annahmeproportion der optimalen Entscheidungsregel beträgt

$$k^* = 0.7432 \ .$$

Wenn also mehr als 74.32% der Befragten starkes Interesse an der neuen Fluglinie bekunden, wird die Nullhypothese verworfen, d.h. Aktion $a_1$ gewählt und die neue Fluglinie eröffnet.

Ein nur geringfügug abweichendes Ergebnis erhält man, wenn die Formel (3.22) für die Bestimmung von k* beim Binomial-Modell verwendet wird. Es ist in diesem Falle

$$r = 148.61$$

und die Annahmeproportion $\frac{r}{n} = 0.7431$.

## 2.25. Ueber den Realitätsgehalt von Entscheidungsmodellen: Kritik der bisher verwendeten Modelle und Vorschläge für deren Ausbau

Alle bisher behandelten Beispiele basieren auf einer Konsequenzen- bzw. Opportunitätskostentabelle mit nur zwei Umweltzuständen. Wegen dieser Beschränkung wirken sie mindestens zum Teil offensichtlich künstlich, denn die Realität, die durch das Entscheidungsmodell abgebildet werden soll, ist ohne Zweifel viel komplexer.

Welches sind nun die Möglichkeiten, die Modelle realer zu gestalten? Ein erster Gedanke mag vielleicht sein, die Anzahl der Umweltzustände zu erhöhen. Beim Fall der PAN ATLANTIC Airways beispielsweise würde die für den Aktor günstige Situation nicht mehr nur durch den Umweltzustand $s_1$ = "Anteil der Interessierten von 80%" repräsentiert, sondern durch $s_1$ = 71%; $s_2$ = 72% ... Die enumerative Lösung eines solchermassen strukturierten Problems lässt sich mit dem hier beschriebenen Verfahren ohne weiteres durchführen. Das "ohne weiteres" bezieht sich allerdings nur auf die methodische Seite des Problems und nicht auf den damit verbundenen rechentechnischen Aufwand, der sofort gewaltig gross wird und der sich nochmals um ein Vielfaches erhöht, wenn die Annahme des konstanten Stichprobenumfanges aufgehoben wird, kurz: Ohne Einbezug eines Computers scheint dieser Weg zur Annäherung an die Realität ein kaum gangbarer zu sein[1].

Es ist auch möglich, auf analytischem Wege Formeln für die Annahmezahl bzw. -Proportion der optimalen Entscheidungsregel für den Fall von mehr als zwei Umweltzuständen zu ermitteln. Der Kompliziertheitsgrad der in den vorgehenden Abschnitten entwickelten Formeln lässt allerdings erahnen, dass auch dieser Weg nicht praktikabel ist, denn die Anzahl der Variablen erhöht sich natürlich mit jedem zusätzlichen Umweltzustand, und zudem erhält man nicht mehr auflösbare exponentielle Gleichungen und müsste die gesuchte Grösse mittels numerischer Methoden näherungsweise

[1] Ein Beispiel für die Ermittlung bedingte Opportunitätskosten-Erwartungswerte für eine Entscheidungsregel (dort "Strategie" genannt) bei 20 Umweltzuständen findet sich bei Chernoff/Moses, S. 195 ff. Andere Beispiele in der Literatur, die etwa auf drei Umweltzuständen aufbauen, sind hingegen kaum als Fortschritt in Richtung realere Modelle zu bezeichnen; vgl. dazu Drukarczyk, S. 4; Topritzhofer, S. 302

ermitteln - bei der Komplexität der Gleichungen ein ebenfalls sehr aufwendiges Verfahren.

Eine weitere Möglichkeit ist das Erfassen der Umweltzustände als eine stetige Variable. Das hat zur Folge, dass die Apriori-Verteilung eine Dichtefunktion ist. Es ist dies der Weg, dessen Beschreibung in der Literatur einen breiten Platz einnimmt und der seinen Niederschlag im bereits erwähnten Konzept der konjugierten Familien von Verteilungen gefunden hat[1]. Der nächste Abschnitt wird sich mit dieser Konzeption befassen.

Eine letzte Möglichkeit der Anpassung der Modelle - es wird im folgenden ersichtlich werden, dass es sich dabei eigentlich eher um eine andere Interpretation der bisher verwendeten Modelle handelt - soll hier kurz skizziert werden.

Bei der Behandlung des Normal-Modells wurden einige Begriffe eingeführt, die aus dem Bereich des Hypothesen-Tests der Statistik stammen. In der Tat ist eine weitgehende Analogie eines 2-Aktionen-Problems mit der Situation des Hypothesen-Tests offensichtlich: Beim Hypothesen-Test wird entschieden, ob die Nullhypothese beibehalten oder zugunsten der Alternativhypothese verworfen werden soll. Bei 2-Aktionen-Problem können die möglichen Umweltzustännde in zwei Bereiche - getrennt durch den Break-even-Punkt - aufgeteilt werden, in einen für den Aktor in bezug auf die zur Verfügung stehenden Aktionen günstigen und einen ungünstigen Bereich. Diese Bereiche lassen sich als Hypothesen auffassen; die Beibehaltung der Hypothese "die Situation ist günstig" bzw. die Verwerfung der Hypothese "die Situation ist ungünstig" haben die Wahl einer bestimmten Aktion zur Folge.

[1] Vgl. S. 50 ff.

Die Menge aller möglicher Umweltzustände $S$ kann also in die beiden sich ausschliessenden Mengen $S_1$ und $S_2$ aufgeteilt werden, und es sind demzufolge die Nullhypothesen

$$H_o : s_w \in S_1 \text{ bzw. } H_o : s_w \in S_2$$

und die entsprechenden Alternativhypothesen

$$H_A : s_w \in S_2 \text{ bzw. } H_A : s_w \in S_1$$

möglich ($s_w$ bezeichnet den wahren Zustand der Welt).

Enthalten $S_1$ und $S_2$ nur je ein Element, so nennt man $s_w \in S_1$ bzw. $s_w \in S_2$ "einfache Hypothesen", sind aber mehr als ein Element in den Zustandsräumen vorhanden, so heissen die Hypothesen "zusammengesetzte Hypothesen". Offenbar wurden in den vorgehenden Abschnitten immer einfache Hypothesen getestet. Diese besitzen die Form einer Gleichung, beim Verpackungsproblem beispielsweise

$$H_o : \mu_1 = 52 \text{ kg},$$

$$H_A : \mu_2 = 50 \text{ kg}.$$

Zusammengesetzte Hypothesen hingegen haben die Form von Ungleichungen, beispielsweise

$$H_o : \mu \geq s_b$$

$$H_A : \mu < s_b$$

mit $s_b$ als dem Break-even-Punkt. Die Hypothesen beim Verpackungsproblem könnten solchermassen lauten:

$$H_o : \mu \geq 51.05 \text{ kg},$$

$$H_A : \mu < 51.05 \text{ kg}.$$

Dementsprechend werden auch die Umweltzustände neu definiert: beim Verpackungsproblem beispielsweise tritt anstelle von $s_1:\mu_1 = 52$ kg $s_1:\mu \geq 51.05$ kg, und $s_2:\mu_2 = 50$ kg wird durch $s_2:\mu < 51.05$ kg ersetzt.

Schliesslich werden durch diese Neuformulierung der Umweltzustände auch die Apriori-Wahrscheinlichkeiten beeinflusst: die Apriori-Wahrscheinlichkeit $p'(s_j)$ des Verpakkungsproblems z.B. ist jetzt die Wahrscheinlichkeit $p(\mu \geq 51.05$ kg) und nicht mehr $p(\mu_1 = 52$ kg). Die Auffassung der Umweltzustände als zusammengesetzte Hypothesen hat zur Folge, dass als Apriori-Verteilung eine stetige, uniforme Verteilung über einem Intervall gewählt werden kann, eine Verteilung, die unter dem Namen "Treppenverteilung" bereits vorgestellt wurde[1]. Die Treppenverteilung wurde aber nicht nur vorgestellt, sondern es wurde auch eine ihrer wesentlichen Eigenschaften bewiesen: Bei Vorliegen einer linearen Konsequenzen- bzw. Opportunitätskostenfunktion und einer Treppenverteilung über dem Zustandsraum, kann der Konsequenzen- bzw. Opportunitätskostenwert der Intervallmitte als Aequivalent für die Funktion innerhalb dieses Intervalls verwendet werden[2].

Die Opportunitätskostenwerte in Tab. 18 können also als die Opportunitätskosten bei $\mu = 50.025$ kg bzw. $\mu = 52.025$ interpretiert werden, die in der Tabelle stellvertretend für eine lineare Opportunitätskostenfunktion stehen.

Diese Interpretation setzt eine Begrenzung der Intervalle nach beiden Seiten voraus, ein Umstand, der keine besonderen Schwierigkeiten verursacht.

---

[1] Vgl. S. 49 f.

[2] Vgl. S. 70 f.

| | $s_1$ : 51.05 kg<μ<54 kg | $s_2$: 49 kg<μ<51.05 kg |
|---|---|---|
| $a_1$ | 0 | 25 |
| $a_2$ | 23 | 0 |
| $p'(s_j)$ | 0.6 | 0.4 |

Tab. 18: Opportunitätskostentabelle mit als zusammengesetzten Hypothesen formulierten Umweltzuständen

Ein Entscheidungsmodell dieses Charakters hat nun u.E. erheblichen Realitätsgehalt, besitzt trotzdem eine einfache Struktur und ist, wie im folgenden ersichtlich wird, den in den vorgehenden Abschnitten entwickelten Lösungsmethoden zugänglich. Dass das Modell eine lineare Konsequenzenfunktion voraussetzt, dürfte in vielen Fällen seiner praktischen Relevanz keinen Abbruch tun.

Im nächsten Abschnitt soll unter dem Titel "Treppenverteilung" zunächst allgemein auf das Testen von Hypothesen eingetreten werden.

Diesem Kapitel folgt eine Betrachtung über die Möglichkeit der Anwendung des Hypothesen-Tests auf das Bayesianische Entscheidungsmodell.

## 2.3. Die Entscheidung bei stetiger Apriori-Verteilung

### 2.31. Treppenverteilung

#### 2.311. Das Testen von Hypothesen in der klassischen Statistik[1]

Im letzten Abschnitt wurde erklärt, warum unter dem Titel "Treppenverteilung" kurz auf den Hypothesen-Test der

[1] Für die Definition des Begriffes "klassische Statistik" ist auf einige Zusammenhänge zu verweisen, die erst im Laufe dieses Abschnittes ersichtlich werden. Die Definition dieses Terminus findet sich deshalb erst am Anfang des nächsten Kapitels.

klassischen Statistik eingegangen werden soll. Die Darstellung beschränkt sich dabei auf das Normal-Modell, zur Illustration wird das Verpackungsproblem verwendet werden.

Beim Testen von Hypothesen spielen folgende vier Parametereine Rolle:

- der α-Fehler;
- der β-Fehler;
- der Stichprobenumfang n;
- die Annahmezahl bzw. -Proportion k, in der Statistik auch "kritischer Wert" genannt.

Beim Test zweier einfacher Hypothesen kann man grundsätzlich zwei dieser vier Parameter frei wählen, die restlichen Parameter werden dadurch determiniert.

Der zentrale Parameter ist für die klassische Statistik der α-Fehler, auch "Signifikanzniveau" des Tests genannt. Es wird empfohlen, zuerst das Signifikanzniveau festzulegen[1], wobei zugestanden wird, dass die Bestimmung der Höhe des α-Fehlers "an sich willkürlich" sei[2].

Dass dem α-Fehler dadurch offensichtlich mehr Gewicht als dem β-Fehler zugemessen wird, erklärt sich durch die bereits erwähnte Vorschrift, dass diejenige Hypothese, welche durch den Test verworfen werden soll, als Nullhypothese ausgezeichnet wird[3]. Die Nullhypothese soll also den "konservativen Standpunkt"[4] vertreten: die Folgen ihrer ungerechtfertigten Verwerfung sind gewichtiger als diejenigen der irrtümlichen Annahme der Alternativhypo-

[1] Hengst, S. 166; Kreyszig, S. 210

[2] Hengst, S. 166

[3] Vgl. S. 109

[4] Kriz, S. 111

these. Mit der primären Festlegung des α-Fehlers - eben des Fehlers I. Art - kann die Wahrscheinlichkeit des Eintritts dieser ungünstigeren Folgen klein gehalten werden.

Was schliesslich den zweiten freien Parameter anbetrifft, so geht die klassische Statistik in vielen Fällen implizit davon aus, dass der Stichprobenumfang gegeben sei[1], bisweilen wird aber auch der Fall, dass nebst dem Signifikanzniveau der β-Fehler festgelegt wird, behandelt[2].

Beim Test zweier zusammengesetzter Hypothesen wird das geschilderte Prozedere dadurch kompliziert, dass bei gegebenem Signifikanzniveau und Stichprobenumfang der kritische Wert k nicht mehr determiniert wird, sondern eine Funktion des wahren Umweltzustandes ist, dasselbe gilt für den zweiten freien Parameter, den β-Fehler. Werden aber der kritische Wert und der Stichprobenumfang gesetzt, so sind sowohl α- als auch β-Fehler Funktionen des wahren Umweltzustandes. Diese Funktionen heissen "Irrtumscharakteristiken"[3]. Die Irrtumscharakteristiken für das Verpackungsproblem sind in Abb. 19 für den Stichprobenumfang n=100 und verschiedene k-Werte dargestellt. Die Nullhypothese sei dabei:

$$H_0 : \mu > 51.05 \text{ kg}$$

Diese Grafik macht auch deutlich, dass bei gegebenem Stichprobenumfang die Verkleinerung des einen Fehlers nur auf Kosten des anderen möglich ist.

Für diesen Fall zweier zusammengesetzter Hypothesen wird in der klassischen Statistik folgendes Verfahren empfohlen:

[1] Beispielsweise Kreyszig, S. 210

[2] Yamane, S. 244 und S. 279 ff.

[3] Vgl. Schlaifer, S. 164

Das Signifikanzniveau beziehe sich auf jenen Umweltzustand im Bereiche der Nullhypothese, bei dem der α-Fehler bei jedem kritischen Wert k am grössten ist. Dies ist, wie aus Abb. 19 ersichtlich ist, gerade der Break-even-Punkt $s_b$ = 51.05 kg. Praktisch heisst das nichts anderes, als dass die zusammengesetzte Nullhypothese durch die

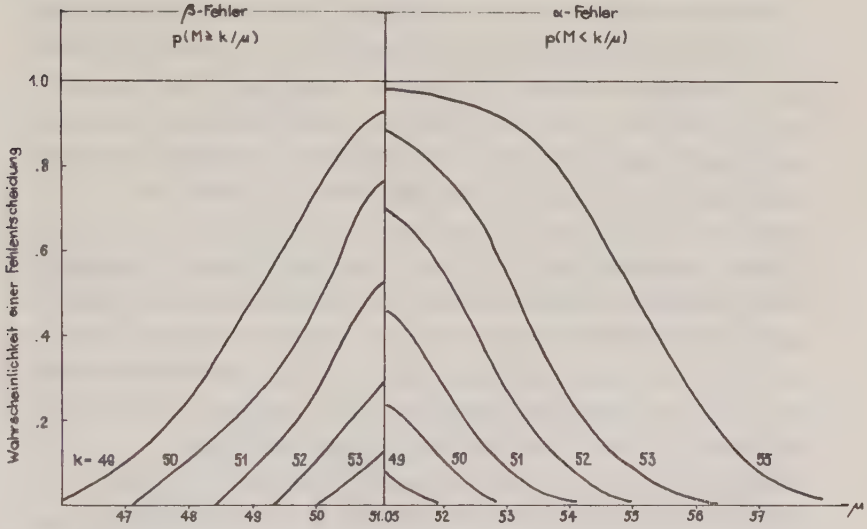


Abb. 19: Irrtumscharakteristiken des Verpackungsproblems (n=100)

einfache Nullhypothese

$$H_0 : \mu = 51.05 \text{ kg}$$

ersetzt wird[1].

[1] Yamane, S. 245

Bei gegebenem Stichprobenumfang ist dadurch k bestimmt. Der β-Fehler wird für den Entscheid im Prinzip gar nicht mehr benötigt, er wird in der Regel bei gegebenem Signifikanzniveau und kritischem Wert als Funktion $\beta(\mu)$ - die Irrtumscharakteristik des β-Fehlers - dargestellt und heisst "Operationscharakteristik", ihr Komplement $1 - \beta(\mu)$ "Gütefunktion"[1].

2.312. Eine "Bayesianische Anwendung" des Hypothesen-Tests

Nachdem das klassische Vorgehen beim Testen von Hypothesen kurz dargestellt wurde, kann endlich der Begriff "klassische Statistik" präzisiert werden. Es wird dieser Terminus vorwiegend von Bayesianern verwendet und dient als Sammelausdruck für all jene Statistik, deren Schlüsse auf einem politischen Entscheid, der Wahl des Signifikanzniveaus, beruhen und die nicht - wie eben beim Bayesianischen Approach - auch von der Höhe der Konsequenzen und der Apriori-Verteilung über dem Zustandsraum beeinflusst werden[2]. Man kann allerdings nicht sagen, dass die derart umschriebene klassische Statistik die Konsequenzen ihrer Schlüsse - besser: die Folgen von Fehlschlüssen - gänzlich ausser Acht lasse, denn die Wahl der Nullhypothese basiert ja auf einem - allerdings nur ordinalen - Vergleich der Folgen der beiden möglichen Fehlschlüsse. Indem aber der Bayesianer die Konsequenzen - und zwar auf einer Ratioskala gemessen - in sein Kalkül miteinbezieht und die optimale Entscheidungsregel aufgrund von Erwartungswerten wählt, setzt er sich dem Einwand der Willkür bei der Festlegung des Signifikanzniveaus nicht mehr aus: er wählt jene $p_\alpha/p_\beta$-Kombination, die den tiefsten

[1] Yamane, S. 250 ff.

[2] Vgl. Moore, S. 298 f.; Winkler, S. 432 ff.

Opportunitätskosten-Erwartungswert zur Folge hat.

Der Bayesianer hat sich aber - um wieder auf das Problem der zusammengesetzten Hypothesen zurückzukommen - mit einem besonderen Problem auseinanderzusetzen: wie werden die Erwartungswerte für die einzelnen Entscheidungsregeln berechnet, wenn $p_\alpha$ und $p_\beta$ keine singulären Werte mehr sind, sondern als Irrtumscharakteristiken vorliegen? Der Gedanke liegt nahe, die heuristische Regel der klassischen Statistik anzuwenden, nämlich den schlechtesten möglichen Umweltzustand sowohl im Bereiche der Nullhypothese wie auch in demjenigen der Alternativhypothese stellvertretend für den ganzen Bereich als einfache Hypothese zu behandeln. Wie gezeigt wurde, braucht der klassische Statistiker diese Regel nur auf den Bereich der Nullhypothese anzuwenden, weil er bei gegebenem Signifikanzniveau den β-Fehler gar nicht kennen muss, um zu einem Entscheid zu gelangen. Der Bayesianer hingegen braucht für seine Wahl der optimalen Entscheidungsregel Kenntnis des β-Fehlers.

Die Ausdehnung der heuristischen Regel auch auf den Bereich der Alternativhypothese hat nun allerdings die absurde Folge, dass Alternativ- und Nullhypothese identisch werden (vgl. Abb. 19).

Wie aus Abb. 19 ersichtlich ist, ist in diesem Falle

$$p_\alpha + p_\beta = 1.$$

Der Opportunitätskosten-Erwartungswert einer Entscheidungsregel $d_k$ ist nach (3.29)

$$Eo(d_k) = p'(s_1)o_{21}p_\alpha + p'(s_2)o_{12}(1-p_\alpha)$$

oder $$Eo(d_k) = p'(s_2)o_{12} + [p'(s_1)o_{21} - p'(s_2)o_{12}]\, p_\alpha.$$

Die Opportunitätskosten-Erwartungswerte sind also eine lineare Funktion des α-Fehlers. Die Steigung dieser Geraden ist positiv, wenn

(1) $p'(s_1)o_{21} > p'(s_2)o_{12}$,

und der Opportunitätskosten-Erwartungswert ist demzufolge minimal, wenn

$$p_\alpha = 0 \text{ bzw. } p_\beta = 0 .$$

Ist aber

(2) $p'(s_1)o_{21} < p'(s_2)o_{12}$

so ist der Opportunitätskosten-Erwartungswert bei

$$p_\alpha = 1 \text{ bzw. } p_\beta = 0$$

minimal.

Wenn Fall (1) gegeben ist, ist die Annahmezahl der optimalen Entscheidungsregel $k=\infty$, bei Fall (2) $k=0$ (vgl. Abb. 19). Dies hat zur Folge, dass auf jeden Fall - und zwar unabhängig vom Resultat der Stichprobe - jene Hypothese mit dem höheren Produkt von Opportunitätskosten und Apriori-Wahrscheinlichkeit verworfen wird. Wie man leicht sieht, ergibt sich auf diese Weise dasselbe Resultat, wie wenn zwischen den alternativen Hypothesen allein aufgrund des Erwartungswertkriteriums (3.2) - also ohne zusätzliche Informationsaktivität - entschieden würde. Der Nutzen zusätzlicher Information wäre unter diesen Umständen gleich Null.

Der "Indifferenzzonen-Approach" von Chernoff/Moses[1] geht von der beschriebenen heuristischen Regel aus, vermeidet aber durch einen Kunstgriff deren absurde Folgen. Ausgangspunkt ist die Ueberlegung, dass der Entscheidende in der Lage sein sollte, in der unmittelbaren Umgebung des Break-

[1] Chernoff/Moses, S. 257

even-Punktes einen Bereich von möglichen Umweltzuständen festzulegen, in dem die Opportunitätskosten vernachlässigt werden können. Der Break-even-Punkt wird auf diese Weise durch einen "Break-even-Bereich", die "Indifferenzzone" ersetzt und die beiden Umweltzustände an den Enden dieser Zone werden als einfache Hypothesen gesetzt. Je grösser die Indifferenzzone gewählt wird, desto kleiner wird der Opportunitätskosten-Erwartungswert der optimalen Entscheidungsregel, weil, wie Abb. 19 zeigt, mit der wachsenden Distanz zwischen den Hypothesen die Risiken von Fehlentscheidungen immer kleiner werden. Es besteht deshalb ein Interesse, die Indifferenzzone möglichst gross zu wählen, andererseits muss man sich dabei aber bewusst bleiben, dass der Indifferenzzonenapproach von der Annahme ausgeht, dass ein Umweltzustand innerhalb der Indifferenzzone bei beiden Aktionen keine Opportunitätskosten zur Folge hat. Diese Prämisse dürfte der Realität umso gerechter werden, je kleiner das Intervall ist. Der Aktor hat bei der Setzung der Indifferenzzone die beiden gegensätzlichen Tendenzen gegeneinander abzuwägen, ein letzter Rest von Willkür kann dabei nicht vermieden werden.

Nach all diesen Ueberlegungen kann folgender Schluss gezogen werden: Die Modelle, die in den Abschnitten 2.21 bis 2.24 behandelt und im Abschnitt 2.25 als wirklichkeitsfremd qualifiziert werden, können doch als adäquate Abbilder der Realität aufgefasst werden. Die beiden Umweltzustände, die als einfache Hypothesen getestet werden, sind die Endpunkte der Indifferenzzone. Sie vertreten die beiden Zustandsräume, die durch die Indifferenzzone getrennt werden. Dies lässt sich mit dem Argument rechtfertigen, dass, wenn der wahre Umweltzustand in einem der beiden Zustandsräume liegt, die Wahrscheinlichkeit eines Fehlentscheides umso grösser wird, je näher dieser Zustand bei der Indifferenzzone liegt und dass es im Interesse des Aktors liegt, mit dem schlechtesten möglichen Fall zu rechnen. Ferner wird angenommen, dass, wenn der wirkliche Umweltzustand im Bereich der Indifferenzzone liegt, es keine Rolle spielt, welche Aktion ergriffen wird.

Die Opportunitätskostenwerte, die den beiden möglichen Fehlentscheiden zugeschrieben werden, sind als Sicherheitsäquivalente einer linearen Opportunitätskostenfunktion innerhalb eines Zustandsraumes zu interpretieren und entsprechen - eine uniforme Apriori-Verteilung über dem Zustandsraum vorausgesetzt - dem Opportunitätskostenwert der Intervallmitte. Daraus folgt schliesslich, dass die Apriori-Wahrscheinlichkeiten, die den beiden Umweltzuständen des Modells zugeschrieben werden, als uniforme Verteilung über den beiden Zustandsräumen aufgefasst werden, also eine Treppenverteilung vorliegt.

Nun bleibt in diesem Abschnitt nichts mehr anderes zu tun, als für die Ermittlung der optimalen Entscheidungsregel bei Vorliegen einer Treppenverteilung auf die in den vorgehenden Abschnitten entwickelten Verfahren zu verweisen.

## 2.32. Konjugierte Familien von Verteilungen

Bei der Einführung des Konzeptes der konjugierten Familien von Verteilungen[1] wurde auf eine wesentliche Eigenschaft dieser Gattung von Verteilungen hingewiesen: Wenn ein Mitglied einer mit einem bestimmten stochastischen Modell konjugierten Familie von Verteilungen als Apriori-Dichte in ein Entscheidungsmodell eingeht und die diesem stochastischen Modell entsprechende Likelihoodfunktion für die Ermittlung der Aposteriori-Dichte verwendet wird, so gehört die Aposteriori-Verteilung der nämlichen Familie von Verteilungen an wie die Apriori-Verteilung. Apriori- und Aposteriori-Verteilung unterscheiden sich nur bezüglich der Werte ihrer Parameter. Dabei gelten die folgen-

[1] Vgl. S. 50 ff.

den Beziehungen[1]:

- Binomial-Modell:

$$\begin{aligned} &\text{Apriori-Verteilung: } f'_{\beta}(\theta;r',n') \\ &\text{Aposteriori-Verteilung: } f''_{\beta}(\theta;r'',n'') \\ &\qquad \text{mit } r'' = r' + r \\ &\qquad\quad n'' = n' + n \end{aligned} \tag{3.32}$$

- Poisson-Modell:

$$\begin{aligned} &\text{Apriori-Verteilung: } f'_{\gamma}(\lambda;\, r',t') \\ &\text{Aposteriori-Verteilung: } f''_{\gamma}(\lambda;\, r'',t'') \\ &\qquad \text{mit } r'' = r' + r \\ &\qquad\quad t'' = t' + r \end{aligned} \tag{3.33}$$

- Normal-Modell

  . $\sigma$ bekannt, $\mu$ unbekannt

$$\begin{aligned} &\text{Apriori-Verteilung: } f'_{N}(\mu;\, m',\, \tfrac{\sigma}{\sqrt{n'}}) \\ &\text{Aposteriori-Verteilung: } f''_{N}(\mu;\, m'',\, \tfrac{\sigma}{\sqrt{n''}}) \\ &\qquad \text{mit } n'' = n' + n \\ &\qquad\quad m'' = \frac{1}{n''}(n'm' + nm) \end{aligned} \tag{3.34}$$

  . $\mu$ bekannt, $\sigma$ unbekannt

$$\begin{aligned} &\text{Apriori-Verteilung: } f'_{\gamma 2}(h;\, \nu',\, v') \\ &\text{Aposteriori-Verteilung: } f''_{\gamma 2}(h;\, \nu'',\, v'') \\ &\qquad \text{mit } \nu'' = \nu' + \nu \\ &\qquad\quad v'' = \frac{1}{\nu''}(\nu' v' + \nu w) \end{aligned} \tag{3.35}$$

[1] Vgl. Raiffa/Schlaifer, S. 53 ff.; ausführliche Herleitungen der hier ohne Beweis stehenden Beziehungen finden sich bei Raiffa/Schlaifer in den Kapiteln 9, 10 und 11, etwas einfacher bei Winkler, S. 157 f. und S. 169 und S. 179

v symbolisiert den Stichprobenumfang, w die Varianz der Stichprobe.

Der folgende - auf der extensiven Methode beruhende - Entscheidungsbaum zeigt die allgemeine Struktur eines Entscheidungsmodells mit einer stetigen Apriori-Verteilung und einer linearen Konsequenzenfunktion[1].

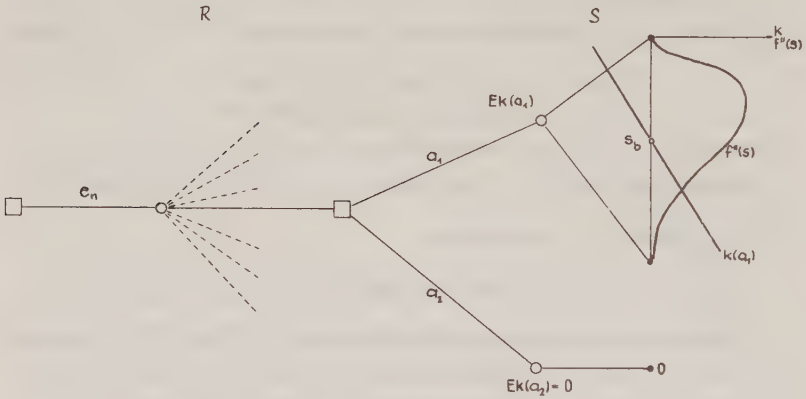


Abb. 20: Entscheidungsbaum für den Fall stetige Apriori-Verteilung/stetige Konsequenzenfunktion

Der Konsequenzen-Erwartungswert $Ek(a_1)$ lässt sich nach (3.5) berechnen, mit einem Unterschied allerdings: hier geht der Aposteriori-Erwartungswert E(S") in die Gleichung ein und nicht E(S').

[1]Der Einfachheit halber sei angenommen, dass $a_1$: "Etwas unternehmen" und $a_2$: "Nichts unternehmen" bedeuten.

Ferner gelten die Relationen

(1) wenn $E(S'') > s_b \rightarrow a_1$;

(2) wenn $E(S'') < s_b \rightarrow a_2$;

(3) wenn $E(S'') = s_b \rightarrow a_1$ oder $a_2$ [1].

Die Aposteriori-Verteilung und damit auch E(S") sind abhängig vom Resultat des Experimentes $e_n$. Wie bei der Herleitung von k* auf dem analytischen Wege bei diskreter Apriori-Verteilung[2], kann auch in diesem Falle die Annahmezahl der optimalen Entscheidungsregel mit Hilfe der Beziehung (3) - dem Punkt, bei dem zwischen den beiden Aktionen nicht unterschieden werden kann - ermittelt werden. k* entspricht demjenigen Resultat, bei dem der Erwartungswert der Aposteriori-Verteilung gerade gleich dem Break-even-Punkt $s_b$ ist.

Es sei dies anhand eines Beispiels mit dem Binomial-Modell erläutert.

In einem Qualitätskontrollfall sei der Break-even-Punkt $s_b = 0{,}1$ und die Apriori-Verteilung eine Betaverteilung $f'_\beta$ (Θ; 1,20) (vgl. Abb. 21). Der Stichprobenumfang betrage n=30.

Der Erwartungswert der Aposteriori-Verteilung ist

$E(S'') = E(\Theta; r'', n'') = \frac{r''}{n''}$ (vgl. Beziehung (2.5 ))

oder nach (3.32)

$E(S'') = E(\Theta; r'', n'') = \frac{r' + r}{n' + n}$ .

[1] Vgl. S. 65 ff.

[2] Vgl. S. 85 ff.

Der Erwartungswert der Aposteriori-Verteilung entspricht dem Break-even-Punkt, wenn

$$0.1 = \frac{r' + r}{n' + n}, \text{ wobei, wie oben erläutert, r gleich der Annahmezahl k* der optimalen Entscheidungsregel ist,}$$

d.h. $0.1 = \frac{1 + k^*}{20 + 30}$

und $k^* = 4$.

$d_4$ ist demzufolge die optimale Entscheidungsregel.

Es sei angenommen, die Stichprobe liefere ein Ergebnis R=5, d.h. der Aktor wird $a_1$ wählen. Die Apriori- bzw. Aposteriori-Verteilung haben in diesem Falle folgendes Aussehen:

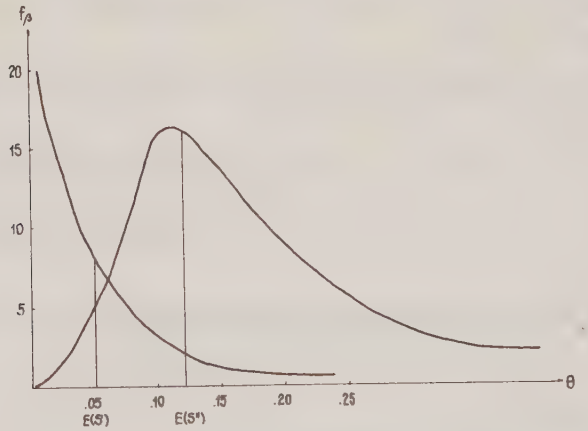


Abb. 21: Konjugierte Apriori- und Aposteriori-Verteilungen

Aufgrund der Apriori-Verteilung ist das erwartete $\Theta=E(S')=$ $= 0.05$. Die Stichprobe ergibt einen Ausschussanteil von $\frac{5}{30} = 0.167$. Die Aposteriori-Verteilung - die mit Hilfe des Bayes-Theorems gewonnene Kombination von Apriori-Wissen mit dem Resultat des Experimentes - lässt ein $\Theta=E(S")=0.12$ erwarten. Dieser Wert liegt zwischen den Werten, die entweder allein aufgrund der Apriori-Verteilung oder allein aufgrund des Stichprobenresultates erwartet werden können.

Zum hier vorgestellten Prozedere gilt es zu bemerken, dass die Kenntnis des Break-even-Punktes ausreicht, um die Annahmezahl der optimalen Entscheidungsregel ermitteln zu können. Es sei dies anhand des Normal-Modells (mit bekanntem $\sigma$ und unbekanntem $\mu$) nochmals demonstriert.

Als Beispiel diene der für diesen Zweck leicht modifizierte Verpackungsfall[1]. Es sind die folgenden Konsequenzenfunktionen gegeben:

$$k(a_1) = -24'500'000 + 480'000\mu$$

$$k(a_2) = 0.$$

Nach (3.6) ist der Break-even-Punkt

$$s_b = \frac{- 24'500'000}{- 480'000} = 51.04 .$$

Apriori-Verteilung sei die mit dem Normal-Modell konjugierte Normal-Verteilung

$$f'_N \left(\mu;\ m'=51,\ \frac{\sigma}{\sqrt{n'}} = 1\right)$$

und der Stichprobenumfang betrage $n=100$.

Weil

$$\sigma = 15,$$

[1] Vgl. S. 97 f.

beträgt

$$n' = 225.$$

k* ist durch die Bedingung

$$E(S'') = E\left(\mu;\ m'', \frac{\sigma}{\sqrt{n''}}\right) = 51.04$$

determiniert.

Da aber

$$E\left(\mu;\ m'', \frac{\sigma}{\sqrt{n''}}\right) = m'',$$

gilt

$$m'' = 51.04$$

Oder nach (3.34)

$$\frac{1}{n''}\,(n'm' + nm) = 51.04$$

bzw.

$$\frac{1}{n'+n}\,(n'm' + nk^*) = 51.04.$$

Mit den Zahlenwerten unseres Beispiels ergibt sich

$$\frac{1}{225 + 100}\,(225 \cdot 51 + 100 \cdot k^*) = 51.04$$

und $k^* = 51.13$.

Auf diese Weise lassen sich die Annahmezahlen der optimalen Entscheidungsregeln auch für die restlichen stochastischen Modelle herleiten.

Die Formeln für die Bestimmung von k* sind im folgenden für alle hier verwendeten stochastischen Modelle zusammengefasst:

- Binomial-Modell: $k^* = (n'+n)s_b - r'$ (3.36)
- Poisson-Modell: $k^* = (t'+t)s_b - r'$ (3.37)

- Normal-Modell:

  . $\sigma$ bekannt, $\mu$ unbekannt: $k^* = \frac{n''}{n}\left(s_b - \frac{n'm'}{n''}\right)$ (3.38)

  . $\mu$ bekannt, $\sigma$ unbekannt: $k^* = \frac{v' + v - s_b\, v'v'}{s_b v}$ (3.39)

## 3. Die Entscheidung mit zusätzlicher Informationsaktivität: variabler Informationsumfang

Nachdem die Methoden der Ermittlung von optimalen Aktionen und Entscheidungsregeln ausführlich behandelt wurden, geht es unter diesem Titel um die Bestimmung optimaler Strategien[1]. Das Entscheidmodell wird um eine Dimension - dem Umfang der zusätzlichen Information - erweitert.

Mit wachsendem Informationsumfang weiss ein Aktor immer mehr über die für ihn relevanten Umweltzustände, die Unsicherheit nimmt also ab. Dies kann in bestimmten Fällen - beispielsweise bei der Beurteilung eines Warenpostens hinsichtlich seines Ausschussanteils - bis zur völligen Beseitigung der Unsicherheit führen, dann nämlich, wenn in unserem Beispiel jedes Element dieses Warenpostens kontrolliert wird.

Dem Streben nach umfassender Information gebieten die dadurch verursachten Kosten Einhalt. Informationen stellen ein wirtschaftliches Gut dar; mit steigendem Informationsumfang steigen die Kosten der beschafften Informationen. Aus dieser Sicht besteht also das Interesse, optimale Entscheide mit möglichst wenig Informationen zu treffen.

[1] Unter einer Strategie wird im folgenden die Kombination eines Experimentes $e_n$ mit der bezüglichen dieses Experimentes optimalen Entscheidungsregel $d^*$ verstanden. Eine solche Strategie wird hier auch $e_n$ - Strategie genannt und wird mit $g(e_n;d^*)$ symbolisiert.

Es gilt nun offenbar, gerade soviel zusätzliche Information einzuholen, dass die Differenz zwischen dem Wert dieser Information und den Kosten ihrer Beschaffung maximiert wird. Der nächste Abschnitt befasst sich mit dem hier umgangssprachlich eingeführten Begriff des Informationswertes.

### 3.1. Der Wert zusätzlicher Information[1]

Der Wert einer Information vom Umfang n wird in der Bayes'schen Entscheidungstheorie als Differenz zwischen dem Konsequenzen-Erwartungswert der $e_n$ - Strategie und dem Konsequenzen-Erwartungswert der optimalen Aktion ohne zusätzliche Informationserhebung bzw. dem Konsequenzen-Erwartungswert der $e_o$ - Strategie definiert[2]. Diese Differenz wird in der deutschsprachigen Literatur "Erwartungswert des Informationsgehaltes einer Stichprobe"[3] oder "erwarteter Wert der Stichprobeninformation"[4], kurz: EIS genannt.

Es gilt also

$$EIS(n) = Ek[g(e_n; d^*)] - Ek[g(e_o; a^*)] \qquad (3.40)$$

$$\text{bzw. } EIS(n) = Eo[g(e_o; a^*)] - Eo[g(e_n; d^*)] \qquad (3.41)$$

Zieht man vom EIS die Kosten der Informationsbeschaffung, c(n), ab, so erhält man den "erwarteten Nettogewinn der

---

[1] Es werden hier ausschliesslich entscheidungsorientierte Informationswert-Begriffe behandelt. Einen Ueberblick über andere Ansätze mit vielen Literaturangaben findet sich bei Lahrmann, S. 76 ff.

[2] Vgl. etwa Lahrmann, S. 85 f.; Weber, S. 77

[3] Topritzhofer, S. 353

[4] Bamberg/Coenenberg, S. 119

Stichprobe vom Umfang n" [1], kurz: ENGS:

$$ENGS(n) = EIS(n) - c(n) \tag{3.42}$$

Ein weiterer in diesem Zusammenhang wesentlicher Informationswert-Begriff ist jener des "erwarteten Wertes vollkommener Information"[2], kurz: EVI. Er ist definiert als Differenz zwischen dem Konsequenzen-Erwartungswert, der bei vollkommener Information (und jeweils nachfolgendem Einsatz der optimalen Aktion) entstehen würde und dem Konsequenzen-Erwartungswert der optimalen Aktion ohne zusätzliche Informationssuche, also der $e_o$-Strategie:

$$EVI = \sum_{j=1}^{J} p'(s_j) \max_i k_{ij} - \max_i \sum_{j=1}^{J} p'(s_j) k_{ij} \tag{3.43}$$

Wie Bamberg/Coenenberg beweisen, ist der EVI aber auch gleich dem Opportunitätskosten-Erwartungswert der $e_o$-Strategie[3], es gilt also auch

$$EVI = Eo\,[g(e_o; a^*)]. \tag{3.44}$$

So ist beispielsweise der EVI für das Verpackungs-Beispiel[4]

$$EVI = 0.6 \cdot 460'000 + 0.4 \cdot 0 - [0.6 \cdot 460'000 + 0.4 \cdot (-500'000)]$$
$$= 200'000.$$

[1] Bamberg/Coenenberg, S. 120

[2] Vgl. Bamberg/Coenenberg, S. 110 f.

[3] Bamberg/Coenenberg, S. 111

[4] Vgl. S. 95 f.

Dieser Wert entspricht, wie man unschwer feststellen kann, dem Opportunitätskosten-Erwartungswert der optimalen Aktion ohne zusätzliche Informationsbeschaffung ($a_1$).

Der EVI kann erstens als in Geldeinheiten gemessener Abstand zu einem Zustande vollkommener Information interpretiert werden oder zweitens als Höchstbetrag, zu dem es sich lohnt, Information zu kaufen. Mit Hilfe des EVI kann denn auch - und das ist seine einzige Funktion - ein für eine bestimmte Situation maximaler Stichprobenumfang $\hat{n}$ ermittelt werden. $\hat{n}$ ist das maximale n, das der Bedingung

$$c(n) \leq EVI \tag{3.45}$$

genügt

und für den optimalen Stichprobenumfang n* gilt

$$0 \leq n^* \leq \hat{n}.$$

Mit Hilfe von (3.44) kann die Beziehung (3.41) noch etwas einfacher geschrieben werden:

$$EIS(n) = EVI - Eo[g(e_n;d^*)]. \tag{3.46}$$

## 3.2. Die optimale Strategie

Die am Anfang dieses Abschnittes umgangssprachlich formulierte Regel, es sei gerade soviel zusätzliche Information einzuholen, dass der Ueberschuss des Wertes dieser Information über die Kosten ihrer Beschaffung möglichst gross sei, kann nun mit Hilfe der eingeführten Informationswert-Begriffe präzis formuliert werden:

Es ist diejenige Strategie zu wählen, für die gilt

$$ENGS(n^*) = \max_n ENGS(n) \tag{3.47}$$

## 3.3. Entscheidung bei diskreter Apriori-Verteilung

### 3.31. Binomial-Modell

Einer exemplarischen Darstellung diene wiederum ein Beispiel aus der Qualitätskontrolle. Es sei folgende Opportunitätskostentabelle gegeben:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 20'000 |
| $a_2$ | 40'000 | 0 |
| $p'(s_j)$ | 0.4 | 0.6 |

Tab. 19: Opportunitätskostentabelle eines Qualitätskontroll-Beispiels

Es seien:

$a_1$ : Warenposten annehmen;
$a_2$ : Warenposten ablehnen;
$s_1$ : Ausschussanteil $\Theta_1$=10%;
$s_2$ : Ausschussanteil $\Theta_2$=20%.

Die Durchführung einer Stichprobe verursacht fixe Kosten von Fr. 1'000.-- und variable Kosten von Fr. 50.-- pro geprüftes Stück, es gilt also

$c = 1'000 + 50\ n.$

Nach (3.44) ist

$E_o[g(e_o;a^*)] = EVI = 12'000$

und nach (3.45)

$\hat{n} = 220.$

Nach (3.47) hat ein Aktor die Strategie mit dem höchsten ENGS zu wählen. Er weiss, dass diese optimale Strategie

eine Stichprobe vom Umfange $0 \leq n^* \leq 220$ beinhaltet. Die Bestimmung der optimalen Strategie geschieht auf dem enumerativen Weg. Dafür sei folgende Heuristik empfohlen: Es sind zunächst die ENGS für drei Strategien zu ermitteln. Die Eintragung dieser Werte in einem Koordinatennetz ermöglicht sodann eine ungefähre Lokalisierung des optimalen Stichprobenumfanges und der genaue Wert kann anschliessend nach wenigen Schritten ermittelt werden[1].

Die Ermittlung der ENGS sei am Beispiel n=20 demonstriert.

In einem ersten Schritt ist die Annahmezahl der optimalen Entscheidungsregel der $e_o$-Strategie zu bestimmen. Nach (3.22) ist

$$r = \frac{\log\left[\frac{40'000}{20'000} \cdot \frac{0.4}{0.6}\right] + 20 \log \frac{0.9}{0.8}}{\log\left[\frac{0.2}{0.1} \cdot \frac{0.9}{0.8}\right]} = 3.26$$

und $k^* = 3$.

Mit Hilfe der Beziehungen (3.20), (3.21), (3.17) und (2.1) kann der Opportunitätskosten-Erwartungswert der $e_o$-Strategie ermittelt werden:

$$E_o[g(e_{20};d_3)] = p'(s_1)o_{21} \sum_{r=4}^{20} p(R=r/\Theta_1) + p'(s_2)o_{12} \sum_{r=0}^{3} p(R=r/\Theta_2)$$

$$= 0.4 \cdot 40'000 \sum_{r=4}^{20} \binom{20}{r} 0.1^r 0.9^{n-r} +$$

$$0.6 \cdot 20'000 \sum_{r=0}^{3} \binom{20}{r} 0.2^r 0.8^{n-r}$$

$$= 16'000 \cdot 0.1329 + 12'000 \cdot 0.4114$$

$$E_o[g(e_{20};d_3)] = 7'063.$$

[1] Dabei ist es eine Frage des Anspruchsniveau, mit welcher Genauigkeit sich ein Aktor zufrieden gibt.

Nach (3.46) ist der

$$EIS(20) = 12'000 - 7'063 = 4'937$$

und

$$c(20) = 1'000 + 50 \cdot 20 = 2'000.$$

Demzufolge ist nach (3.42)

$$ENGS(20) = 4'937 - 2'000 = 2'937.$$

Diese Werte und die analog ermittelten Resultate für weitere Strategien sind in Tab. 20 zusammengefasst. Der Stichprobenumfang der optimalen Strategie - der $e_{70}$-Strategie - beträgt danach $n^* = 70$.

Es wird aus dieser Tabelle ersichtlich, dass der Opportunitätskosten-Erwartungswert einer Strategie mit zunehmendem Stichprobenumfang immer kleiner wird. Es ist dies eine Folge davon, dass mit wachsendem n die Summe der Wahrscheinlichkeiten, mit denen die Oppportunitätskosten gewichtet werden, abnimmt oder mit anderen Worten das Risiko von Fehlentscheidungen mit zunehmendem Informationsstand kleiner wird.

| n | k* | $E_0[g(e_n; d^*)]$ | EIS(n) | c(n) | ENGS(n) |
|---|---|---|---|---|---|
| 20 | 3 | 7'063 | 4'937 | 2'000 | 2'937 |
| 50 | 7 | 4'238 | 7'762 | 3'500 | 4'262 |
| 68 | 10 | 3'276 | 8'724 | 4'400 | 4'324 |
| 69 | 10 | 3'219 | 8'781 | 4'450 | 4'331 |
| 70 | 10 | 3'157 | 8'843 | 4'500 | 4'343 |
| 71 | 10 | 3'125 | 8'875 | 4'550 | 4'325 |
| 100 | 14 | 2'126 | 9'874 | 6'000 | 3'874 |

Tab. 20: Die ENGS(n) für das Qualitätskontroll-Beispiel

Der Zusammenhang zwischen den in Tab. 20 aufgeführten Grössen wird aus der Abb. 22 ersichtlich.

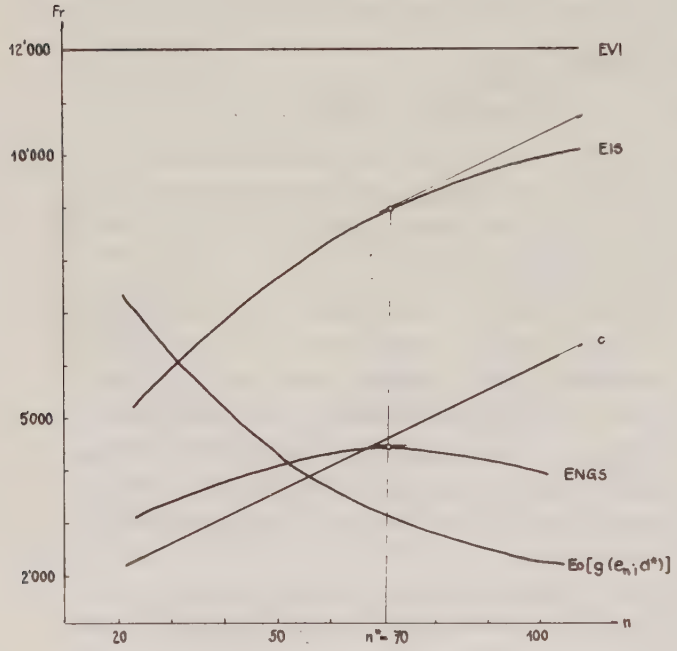


Abb. 22: EVI, EIS(n), c(n), ENGS(n) und Eo(n) für das Qualitätskontroll-Beispiel

Aus der Abbildung wird deutlich, dass sich die EIS asymptotisch dem EVI nähern und dass es sich bei der Kurve der EIS um eine Spiegelung der Eo-Kurve an einer horizontalen Symmetrieachse handelt. Asymptote der Eo-Kurve ist die Abszisse. Die Bedingung (3.47) ist bei jenem Abszissenwert erfüllt, bei dem die Steigungen der EIS-Kurve und der Kostengeraden gleich gross sind.

## 3.32. Hypergeometrisches Modell

Eine Flugzeugfirma steht vor folgendem Problem:

Eine Abteilung der Unternehmung produziert jeweils 50 Stücke eines wichtigen, tragenden Flugzeugteils. Diese Stücke können wohl optisch auf ihre Belastungsfähigkeit hin untersucht werden, sicheren Aufschluss darüber, ob ein bestimmtes Stück die geforderte Belastungsfähigkeit erträgt, bringt aber nur eine das Stück zerstörende Kontrolle. Soll die Firma eine derartige Kontrolle eines Teils der Produktion durchführen, wenn ja, wie gross soll der Stichprobenumfang sein?

Die Konsequenzentabelle für dieses Problem hat das folgende Aussehen:

| | $s_1$ | $s_2$ | $s_3$ |
|---|---|---|---|
| $a_1$ | 0 | -20'000'000 | -40'000'000 |
| $a_2$ | -300'000 | -300'000 | -300'000 |
| $p'(s_j)$ | 0.99 | 0.009 | 0.001 |

Tab. 21: Konsequenzentabelle für das Flugzeug-Beispiel

Dabei bedeuten:

$a_1$ : Der Posten wird für die Montage verwendet;

$a_2$ : Der Posten wird nicht für die Montage verwendet;

$s_1$ : es sind alle Stücke des Postens der geforderten Belastung gewachsen: $\theta_1 = 0$;

$s_2$ : ein Stück des Postens erbringt die geforderte Belastungsfähigkeit nicht: $\theta_2 = 0.02$;

$s_3$ : zwei Stücke sind der Belastung nicht gewachsen, $\theta_3 = 0.04$.

Nach (3.43) ist

$$EVI = 0.99 \cdot 0 + 0.009(-300'000) + 0.001(-3,00'000) - [0.009 \cdot (-20'000'000) + 0.001(-40'000'000)]$$
$$= 217'000.$$

Für die Kosten der Stichprobe gelte

$$c = 4'000\ n + 50\ n^2$$

und nach (3.45)

$$\hat{n} = 37.$$

Die Ermittlung der ENGS sei am Beispiel n=10 demonstriert. Dazu sei das Problem zunächst mit Hilfe eines Entscheidbaumes - nach der extensiven Form der Analyse - dargestellt.

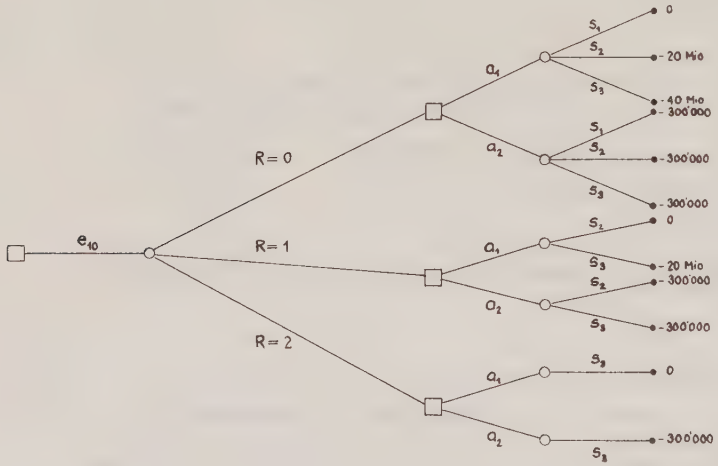


Abb. 23: Entscheidungsbaum für das Flugzeug-Beispiel

Der Konsequenzen- bzw. Opportunitätskosten-Erwartungswert einer $e_n$-Strategie wird bei der extensiven Form der Analyse mit Hilfe der sogenannten Rückwärtsanalyse er-

mittelt[1]. Ein Blick auf Abb. 23 zeigt, dass es in unserem Beispiel vorteilhafter ist, mit den Konsequenzen als mit den Opportunitätskosten zu rechnen, weil die Konsequenzen der Alternative $a_2$ sicher und zudem die Konsequenzen $k(a_1/s_1)$ gleich Null sind.

Allgemein gilt

$$(1) \quad Ek[g(e_{10};d^*)] = \sum_{r=0}^{2} p(R=r) \max_i \sum_{j=r+1}^{3} k_{ij}\, p''(s_j/R=r).$$

Nach dem Bayes-Theorem (1.2) ist aber

$$p''(s_j/R=r) = \frac{p'(s_j)p(R=r/s_j)}{\sum_{j=1}^{J} p'(s_j)p(R=r/s_j)}$$

und nach (1.1)

$$p(R=r) = \sum_{j=1}^{J} p'(s_j)p(R=r/s_j).$$

Eingesetzt in (1) erhält man

$$Ek[g(e_{10};d^*)] = \sum_{r=0}^{2} p(R=r) \max_i \sum_{j=r+1}^{3} k_{ij} \frac{p'(s_j)p(R=r/s_j)}{p(R=r)}$$

und durch Kürzen von $p(R=r)$

$$Ek[g(e_{10};d^*)] = \sum_{r=0}^{2} \max_i \sum_{j=r+1}^{3} k_{ij}\, p'(s_j)p(R=r/s_j).$$

[1] Vgl. S. 78

Dieses Beispiel zeigt, dass das Bayes-Theorem auch bei der extensiven Form der Analyse nicht angewendet werden muss, weil die totalen Wahrscheinlichkeiten p(R=r) gekürzt werden können[1].

Tab. 22 zeigt ein Rechenschema, mit dessen Hilfe der Konsequenzen-Erwartungswert der $e_{10}$-Strategie ermittelt werden kann.

| r | i | j | $p(R=r/s_j)$ | $\sum_{j=r+1}^{3} k_{ij} p'(s_j) p(R=r/s_j)$ | $\max_i \sum_{j=r+1}^{3} k_{ij} p'(s_j) p(R=r/s_j)$ |
|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) |
| 0 | 1 | 1 | - | | |
| 0 | 1 | 2 | .8 | | |
| 0 | 1 | 3 | .6367 | -169'469 | -169'469 |
| 0 | 2 | 1 | 1 | | |
| 0 | 2 | 2 | .8 | | |
| 0 | 2 | 3 | .6367 | -299'351 | |
| 1 | 1 | 2 | - | | |
| 1 | 1 | 3 | .3265 | -6'530 | |
| 1 | 2 | 2 | .2 | | |
| 1 | 2 | 3 | .3265 | -638 | -638 |
| 2 | 1 | 3 | - | 0 | |
| 2 | 2 | 3 | - | <0 | 0 |
| | | | | $\sum_{r=0}^{2}$ | -170'107 |

Tab. 22: Rechenschema für die Ermittlung des Konsequenzen-Erwartungswertes der $e_{10}$-Strategie für das Flugzeug-Beispiel

[1] Vgl. dazu auch Altrogge, S. 827 f.

Erläuterungen zum Rechenschema:

- Spalte (4)

Die Likelihoods $p(R=r/s_j)$ werden mit Hilfe des Hypergeometrischen Modells (Beziehung (2.2)) ermittelt.

Für r = 0, i = 1 und j = 2 gilt beispielsweise

$$p(R=0;n=10,\theta_2=0.02) = \frac{\binom{1}{0}\binom{49}{10}}{\binom{50}{10}} = 0.8 \ .$$

Ein Strich in Spalte (4) bedeutet, dass die Likelihood dieser r/i/j-Kombination - wie aus dem Entscheidungsbaum ersichtlich - für die Ermittlung des Konsequenzen-Erwartungswertes nicht benötigt wird.

- Spalte (5)

Für r = 0 und i = 1 ist beispielsweise

$$\sum_{j=r+1}^{3} k_{ij}\, p'(s_j)\, p(R=r/s_j) = 0 \cdot 0.99 \cdot x + (-20'000'000) \cdot 0.009 \cdot 0.8 + (-40'000'000) \cdot 0.001 \cdot 0.6367 = -169'468 \ .$$

- Spalte (6)

Aus dieser Spalte lässt sich die optimale Entscheidungsregel für die $e_{10}$-Strategie ablesen.

Sie lautet:

wenn $R = 0 \rightarrow a_1$;

wenn $R = 1 \rightarrow a_2$;

wenn $R = 2 \rightarrow a_1$.

Das Stichprobenresultat R = 2 führt in diesem Beispiel zur Wahl von $a_1$, weil der Aktor bei zwei schlechten Stükken in der Stichprobe mit Sicherheit weiss, dass die nicht geprüften Stücke den Anforderungen genügen.

Weil

$$Ek[g(e_0; a^* = a_1)] = 0.009(-20'000'000)+0.001(-40'000'000)$$
$$= - 220'000,$$

gilt nach (3.40)

$$EIS(10) = -170'107 + 220'000 = 49'893 .$$

Tab. 23 fasst einige EIS(n) und ENGS(n) zusammen. Es wird dabei ersichtlich, dass die $e_9$-Strategie die optimale Strategie ist. (Die Entscheidungsregel bleibt für alle Stichprobenumfänge dieselbe.)

| n | EIS(n) | c(n) | ENGS(n) |
|---|---|---|---|
| 5 | 25'349 | 21'250 | 4'099 |
| 8 | 40'182 | 35'200 | 4'982 |
| 9 | 45'048 | 40'050 | 4'998 |
| 10 | 49'893 | 45'000 | 4'893 |
| 15 | 73'632 | 71'250 | 2'382 |

Tab. 23: Die ENGS(n) für das Flugzeug-Beispiel

### 3.33 Poisson-Modell

Eine Unternehmung der Maschinenindustrie habe folgendes Problem zu lösen:

Die Entwicklungsabteilung der Firma hat ein Projekt für eine neue Maschine zur Ausführungsreife gebracht. In der Marketingabteilung sind die Meinungen über die Chancen der neuen Maschine auf dem Markt kontrovers. Man beschliesst, die Frage zu überprüfen, ob allenfalls ein versuchsweiser

Bau und Verkauf der betreffenden Maschine in Frage käme und wenn ja, wie lange diese Versuchsperiode zu bemessen wäre.

Der mit der Lösung dieses Problems betraute Stabsmitarbeiter stellt die folgende Konsequenzentabelle zusammen:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 1'000'000 | -4'000'000 |
| $a_2$ | -1'000'000 | -1'000'000 |
| $p'(s_j)$ | 0.4 | 0.6 |

Tab. 24: Konsequenzentabelle für das Maschinen-Beispiel

Dabei bedeuten:

$a_1$ : Projekt ausführen;

$a_2$ : Projekt nicht ausführen;

$s_1$ : Die Marktlage ist günstig: Während fünf Jahren werden pro Jahr 50 Maschinen verkauft oder pro Tag $\lambda_1 = 0.2$ Maschinen.

$s_2$ : Die Marktlage ist ungünstig: Es werden im gleichen Zeitraum bloss 25 Maschinen pro Jahr oder $\lambda_2 = 0.1$ pro Tag abgesetzt.

Bei den Konsequenzenwerten $k_{1j}$ handelt es sich um Kapitalwerte bei Vorliegen der Situationen $s_1$ bzw. $s_2$, bei den Werten $k_{2j}$ um die Entwicklungskosten. Die Stichprobe bezieht sich hier - wie beim Autoverkäufer-Beispiel[1] - auf ein Zeitintervall t. Dabei gelte die folgende Annahme: Die Kapazität der Firma ist unbegrenzt; es können also beliebig

[1] Vgl. S. 92 f.

viele dieser neuen Maschinen hergestellt und verkauft werden.

Die Kosten der Stichprobe setzen sich aus den folgenden Komponenten zusammen:

- Dem Fixkostenbestandteil f = 50'000, der durch das Umrüsten einer Abteilung verursacht wird.
- Einem entgangenen Deckungsbeitrag pro Tag x = 15'000, der dadurch verursacht wird, dass das alte Produkt, das in dieser Abteilung gefertigt wurde, während der Probezeit nicht mehr hergestellt werden kann. Es sei dabei angenommen, dass dieses Produkt ohne weiteres wieder hergestellt und im gleichen Umfange wie in der Vergangenheit wieder verkauft werden könnte.
- Den Kosten der Umrüstung im Falle eines negativen Ergebnisses der Versuchsperiode y = 30'000. Weil dieses Ergebnis unsicher ist - eine Unsicherheit, die durch die Apriori-Wahrscheinlichkeit gemessen wird - handelt es sich hier um die erwarteten Kosten $p'(s_2) \cdot y$ oder $0.6 \cdot 30'000 = 18'000$.

Die Summe dieser Kostenkomponenten reduziert sich um den Deckungsbeitrag, den die während der Versuchsperiode verkauften Maschinen einbringen. Der Deckungsbeitrag pro verkaufte Maschine beträgt z = 80'000, der gesamte Dekkungsbeitrag ist gleich dem Produkt von Anzahl verkaufter Maschinen r mit diesem Deckungsbeitrag z. Weil es sich bei dieser Anzahl verkaufter Maschinen um eine Zufallsvariable handelt, ist auch dieser konstenmindernde Bestandteil ein Erwartungswert.

Die erwarteten Kosten sind

$$Ec = f + tx + p'(s_2)y - [p'(s_1)\ E(R/\lambda_1 t) \cdot z + p'(s_2) E(R/\lambda_2 t) z]$$

Als für die hier geschilderte Problemlage adäquates stochastisches Modell bietet sich das Poisson-Modell an, und es gilt deshalb

$$E\,(R/\lambda_j t) = \lambda_j t$$

und

$$Ec = f + tx + p'(s_2)y - [p'(s_1)\lambda_1 t\, z + p'(s_2)\lambda_2 t \cdot z]$$

bzw.

$$Ec = f + p'(s_2)y + t\,[x - (p'(s_1)\cdot\lambda_1 + p'(s_2)\lambda_2)\, z].$$

Mit den Werten unseres Beispiels ergibt sich:

$$Ec = 50'000 + 18'000 + t\,[15'000 - (0.4\cdot 0.2 + 0.6\cdot 0.1)80'000]$$

$$Ec = 68'000 + 3'800\ t\ .$$

Zunächst sei die Konsequenzentabelle in eine Opportunitätskostentabelle transformiert:

| | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 3'000'000 |
| $a_2$ | 2'000'000 | 0 |
| $p'(s_j)$ | 0.4 | 0.6 |

Tab. 25: Opportunitätskostentabelle für das Maschinen-Beispiel

Nach (3.44) ist

$$Eo\,[g(e_o;a^*)] = EVI = 800'000$$

und nach (3.45)

$$\hat{t} = 192.$$

Die Ermittlung der ENGS sei wiederum exemplarisch - für n = 40 - demonstriert.

Die Annahmezahl der optimalen Entscheidungsregel für $e_{40}$ wird mit Hilfe von Beziehung (3.26) bestimmt. Es ist

$$r = \frac{\ln \frac{2}{3} \cdot \frac{0.4}{0.6} + t\ (0.1 - 0.2)}{\ln (\frac{0.1}{0.2})} = 6.94$$

und

$k^* = 7.$

Aufgrund von (3.24), (3.25), (3.17) und (2.4) gilt

$$E_0 [g(e_{40};d_7)] = p'(s_1)\ o_{21} \sum_{r=0}^{6} p(R=r/s_1) + p'(s_2)\ o_{12} \sum_{r=7}^{\infty} p(R=r/s_2)$$

$$= 0.4 \cdot 2'000'000 \cdot \sum_{r=0}^{6} \frac{8^r}{r!} e^{-8} + 0.6 \cdot 3'000'000 \cdot \sum_{r=7}^{\infty} \frac{4}{r!} e^{-4}$$

$$= 800'000 \cdot 0.3133 + 1'800'000 \cdot 0.1106$$

$$= 449'720.$$

Nach (3.43) ist der

EIS (40) = 800'000 - 449'720 = 350'280.

Die Kosten der Informationsbeschaffung belaufen sich auf

c(40) = 68'000 + 40·3'800 = 220'000.

Nach (3.42) ist der

ENGS (40) = 350'280 - 220'000 = 130'280.

Die analog ermittelten Resultate für weitere Strategien sind in Tab. 26 zusammengefasst. Die optimale Strategie lautet:

Die projektierte Maschine ist während 65 Tagen probeweise herzustellen und zu verkaufen. Werden während dieser Periode sieben oder mehr Maschinen verkauft, so ist das Projekt definitiv durchzuführen.

| t | k* | Eo $g(e_t;d^*)$ | EIS (t) | Ec (t) | ENGS (t) |
|---|---|---|---|---|---|
| 40 | 7 | 449'720 | 350'280 | 220'000 | 130'280 |
| 60 | 10 | 344'760 | 455'240 | 296'000 | 159'240 |
| 63 | 11 | 329'400 | 470'600 | 307'400 | 163'200 |
| 65 | 11 | 321'600 | 478'400 | 315'000 | 163'400 |
| 70 | 12 | 303'940 | 496'060 | 334'000 | 162'060 |
| 80 | 13 | 269'140 | 530'860 | 372'000 | 158'860 |

Tab. 26: Die ENGS(t) für das Maschninen-Beispiel

## 3.34. Normal-Modell

Die Herleitung der otpimalen Strategie unter Verwendung des Normal-Modells sei am Verpackungsbeispiel gezeigt[1].

Es seien nochmals die Parameter dieses Falls in Erinnerung gerufen:

$\sigma_{21}$ = 460'000 $\mu_1$ = 52 kg $p'(s_1)$ = 0.6 $\sigma^2$ = 225 kg

$\sigma_{12}$ = 500'000 $\mu_2$ = 50 kg $p'(s_2)$ = 0.4

Für die Kosten der Stichprobe gelte:

$$c = 20'000 + 500\ n.$$

[1] Vgl. S. 97 f.

Der EVI ist

$$EVI = 200'000$$

und

$$\hat{n} = 360.$$

Der ENGS für n=50 berechnet sich wie folgt:

Zunächst muss auch in diesem Falle die Annahmezahl der optimalen Entscheidungsregel für $e_{50}$ ermittelt werden. Aufgrund von Beziehung (3.30) und den Werten unseres Beispieles ist

$$k^* = \frac{n \cdot 0.90667 - 0.64417}{n \cdot 0.01778}$$

$$= \frac{50 \cdot 0.90667 - 0.64417}{50 \cdot 0.01778} = 50.275 .$$

Nach (3.29) ist

$$E_0 [g(e_{50}; d_{50.275})] = 0.6 \cdot 460'000 \cdot p_\alpha + 0.4 \cdot 500'000 \cdot p_\beta .$$

Für $z_\alpha$, mit dessen Hilfe sich aus einer Tabelle für die standardisierte Normalverteilung der Wert für $p_\alpha$ ablesen lässt, gilt

$$z_\alpha = \frac{k - \mu_1}{\sigma_m} \quad \text{und} \quad \sigma_m = \frac{n}{\sqrt{n}}$$

oder

$$z_\alpha = \frac{50.275 - 52}{\frac{15}{\sqrt{50}}} = - 0.8132 .$$

Demnach ist

$$p_\alpha = 0.2081 \; .^{1}$$

Analog ist

$$z_\beta = \frac{k-\mu_2}{\sigma_m} = \frac{50.275 - 50}{\frac{15}{\sqrt{50}}} = 0.1296$$

und

$$p_\beta = 0.4486.$$

Eingesetzt in (3.29) ergibt sich

$$E_0 [g(e_{50};d_{50.275})] = 0.6 \cdot 460'000 \cdot 0.2081 + 0.4 \cdot 500'000 \cdot 0.4486$$

$$= 147'156.$$

Nach (3.46) ist

$$EIS(50) = 200'000 - 147'156 = 52'844.$$

Die Kosten dieser Stichprobe belaufen sich auf

$$c = 20'000 + 50 \cdot 500 = 45'000$$

und es ist

$$ENGS(50) = 52'844 - 45'000 = 7'844.$$

---

[1] Es empfielt sich, die z-Werte nicht auf die in den Tabellen üblichen zwei Stellen nach dem Komma zu runden, sondern zu interpolieren, weil andernfalls die ENGS-Kurve Sprünge enthält.

Tab. 27 fasst die analog ermittelten Werte für andere Strategien zusammen. Optimale Strategie ist die $e_{110}$-Strategie mit k* = 50.671 kg.

| n | k* | $Eo[g(e_n;d^*)]$ | EIS(n) | c(n) | ENGS(n) |
|---|---|---|---|---|---|
| 50 | 50.275 | 147'156 | 52'844 | 45'000 | 7'844 |
| 100 | 50.638 | 117'810 | 82'190 | 70'000 | 12'190 |
| 110 | 50.671 | 112'586 | 87'414 | 75'000 | 12'414 |
| 120 | 50.698 | 108'168 | 91'832 | 80'000 | 11'832 |
| 200 | 50.819 | 80'641 | 119'359 | 120'000 | -641 |

Tab. 27: Die ENGS(n) für das Verpackungs-Beispiel

## 3.4. Entscheidung bei stetiger Apriori-Verteilung

### 3.41. Treppenverteilung

Unter diesem Titel bleibt nichts anderes übrig, als auf das Kapitel 2.31 dieses Teils zu verweisen, in dem das Entscheidungsmodell mit einer Treppenverteilung als Apriori-Verteilung und einer linearen Konsequenzenfunktion zurückgeführt wird auf den bereits behandelten Fall von Entscheidungsmodellen mit zwei Zustandsvariablen, einer diskreten Apriori-Verteilung und diskreter Konsequenzenfunktion. Für die Ermittlung der optimalen Strategie sind deshalb die im Abschnitt 3.3 dieses Teils vorgestellten Methoden anzuwenden.

### 3.42. Konjugierte Familien von Verteilungen

Dieses Kapitel über die Ermittlung optimaler Strategien bei Vorliegen einer mit einem stochastischen Modell konjugierten Apriori-Verteilung beschränkt sich auf die Darstellung einiger wesentlicher Resultate aus der Literatur. Massgebliche Quelle für dieses Teilgebiet Bayesianischer

Entscheidungstheorie ist das bereits mehrfach zitierte Werk von Raiffa/Schlaifer. Für Herleitungen und Beweise wird auf dieses Buch verwiesen.

In diesen Ausführungen wird die folgende Notation verwendet: $p(R)$ sei die Massefunktion einer diskret verteilten Zufallsvariablen und es ist

$$F(a) = \sum_{R=-\infty}^{a} p(R)$$

$$\text{bzw. } G(a) = \sum_{R=1}^{\infty} p(R)$$

Es gilt also in diesem Fall

$$F(a) + G(a) = 1 + p(R=a).$$

Bei Vorliegen der Dichtefunktion $f(x)$ ist

$$F(a) = \int_{-\infty}^{a} f(x)\, dx$$

bzw.

$$G(a) = \int_{a}^{\infty} f(x)\, dx.$$

Hier ist also

$$F(a) + G(a) = 1.$$

Ein Index verweist jeweils auf die zugrundeliegende Masse- bzw. Dichtefunktion. Neben den bereits bekannten $\beta, \gamma$ und N werden als Indices verwendet:

bin für eine binomial-verteilte Zufallsvariable;
P für eine poisson-verteilte Zufallsvariable und
nbin für eine negativ-binomial-verteilte Zufallsvariable.

## 3.421. Allgemeine Theorie

### 3.421.1. Die Ermittlung des EVI

Nach (3.44) entspricht der EVI den erwarteten Opportunitätskosten der optimalen Aktion vor zusätzlicher Informationsbeschaffung. Im Falle stetiger Apriori-Verteilung und linearer Konsequenzenfunktion kann dieser Wert aufgrund von Beziehung (3.13) bzw. (3.14) ermittelt werden.

Es seien nochmals die beiden in diesen Gleichungen vorkommenden Integrale in Erinnerung gerufen:

$$\int_{s_b}^{\infty} (s-s_b)\, f'(s)\, ds \quad \text{bzw.} \quad \int_{-\infty}^{s_b} (s_b-s)\, f'(s)\, ds \quad [1].$$

Diese Integrale werden "rechtes" bzw. "linkes lineares Loss-Integral"[2] genannt. Das rechte bzw. linke lineare Loss-Integral für irgendeine Zufallsvariable x und die Dichtefunktion f(x) ist wie folgt definiert:

$$L_x^r(a) = \int_a^{\infty} (x-a) f(x)\, dx \tag{3.48}$$

bzw.

$$L_x^l(a) = \int_{-\infty}^{a} (a-x) f(x)\, dx. \tag{3.49}$$

und es gilt

$$EVI = \begin{cases} |t_2-t_1|\, L_{s'}^r(s_b) & \text{für } E(S') \leq s_b \\ |t_2-t_1|\, L_{s'}^l(s_b) & \text{für } E(S') > s_b. \end{cases} \tag{3.50}$$

[3]

---

[1] Vgl. S. 68

[2] Aus Gründen der sprachlichen Einfachheit sei hier das englische "Loss" dem deutschen "Opportunitätskosten" vorgezogen.

[3] Raiffa/schlaifer, S. 97; Winkler, S. 361

### 3.421.2. Die Ermittlung des EIS

Raiffa/Schlaifer zeigen, dass zwischen dem EVI und EIS eine formale Analogie besteht, es gilt nämlich

$$EIS(n) = \begin{cases} |t_2 - t_1| \, L^{r}_{E(S'')}(s_b) & \text{für } E(S'') \leq s_b \\ |t_2 - t_1| \, L^{l}_{E(S'')}(s_b) & \text{für } E(S'') > s_b \end{cases} \quad (3.51)$$

Der einzige Unterschied zum EVI ist, dass sich die linearen Loss-Integrale auf eine andere Zufallsvariable, nämlich auf den Aposteriori-Erwartungswert, bezieht. Die Funktion dieser Zufallsvariablen $f(E(S'')/e_n)$ heisst "Apriori-Verteilung des Aposteriori-Erwartungswertes" oder "Präposteriori-Verteilung"[2].

### 3.421.3. Die Ermittlung der optimalen Strategie

Mit Hilfe des EIS(n) und der Stichprobenkosten kann n*, der Stichprobenumfang der optimalen Strategie, aufgrund der Beziehung (3.47) ohne weiteres ermittelt werden. Die bezüglich des Experimentes $e_{n*}$ optimale Entscheidregel lässt sich - je nach dem verwendeten stochastischen Modell - mittels der Beziehungen (3.36) - (3.39) herleiten.

[1] Raiffa/Schlaifer, S. 104 f.

[2] A.a.O. Winkler spricht in diesem Zusammenhang von der "predictive distribution", Winkler, S. 113 ff. und S. 204 ff.

## 3.422. Normal-Modell[1]

### 3.422.1. Die Ermittlung des EVI[2]

Bei Vorliegen der Normalverteilung

$f'_N(\mu; m', \sigma')$

als Apriori-Verteilung, gilt für das rechte bzw. linke lineare Loss-Integral:

$$L^r_\mu(s_b) = \int_{s_b}^{\infty} (\mu - s_b) f'_N(\mu; m', \sigma') d\mu = \sigma' \, L_{N*}(D) \tag{3.52}$$

$$L^l_\mu(s_b) = \int_{-\infty}^{s_b} (s_b - \mu) f'_N(\mu; m', \sigma') d\mu = \sigma' \, L_{N*}(-D) \tag{3.53}$$

mit

$$D = \frac{s_b - m'}{\sigma'}. \tag{3.54}$$

$L_{N*}(D)$ ist das lineare Loss-Integral für die standardisierte Normalverteilung und ist folgendermassen definiert:

$$L_{N*}(D) = \int_{D}^{\infty} (z - D) \, f_N(z) \, dz \tag{3.55}$$

Die Werte für $L_{N*}(D)$ sind tabelliert[3].

---

[1] Aus einem didaktischen Interesse wird hier von der bisher gehandhabten Reihenfolge bei der Behandlung der stochastischen Modelle abgegangen und zuerst das Normal-Modell, das in diesem Kontext am wenigsten Schwierigkeiten bereitet, besprochen. Im übrigen beschränken wir uns auch hier auf den Fall mit unbekanntem $\sigma$ und unbekannten $\mu$.

[2] Raiffa/Schlaifer, S. 295; Winkler, S. 361 ff.

[3] Raiffa/Schlaifer, S. 356; Winkler, S. 516 ff.

Es gilt nach (3.50)

$$EVI = \begin{cases} |t_2-t_1|\ \sigma'\ L_{N*}(D) & \text{für } m' \leq s_b \\ |t_2-t_1|\ \sigma'\ L_{N*}(-D) & \text{für } m' > s_b\,. \end{cases} \tag{3.56}$$

Die Ermittlung des EVI sei am Verpackungs-Beispiel demonstriert. Es sind

$s_b = 51.04;\ t_1 = 480'000;\ t_2 = 0$

und

$f'(s) = f'_N(\mu; m'=51, \sigma'=1)$.

Nach (3.56) ist

$EVI = 480'000 \cdot 1 \cdot L_{N*}(D)$

und nach (3.54)

$$D = \frac{51.04 - 51}{1} = 0.04.$$

Aus der Tabelle lässt sich ablesen

$L_{N*}(0.04) = 0.3793.$

Man erhält also

$EVI = 480'000 \cdot 0.3793 = 182'064.$

3.422.2. Die Ermittlung des EIS

Wenn die Apriori-Verteilung eine Normalverteilung ist und als stochastisches Modell das Normal-Modell angewendet wird, so handelt es sich bei der Apriori-Verteilung des Aposteriori-Erwartungswertes wiederum um eine Normal-Verteilung mit dem Erwartungswert

$$E\,(E(S");m',n';n) = E(S') = m' \tag{3.57}$$

und der Varianz

$$V\,(E(S");m',n';n) = \frac{n\sigma^2}{(n'+n)n'} = \frac{n\sigma'^4}{\sigma^2+n\sigma'^2} = \sigma^{*2}.^{1} \tag{3.58}$$

Das Problem der Ermittlung des EIS ist nun formal genau das gleiche wie dasjenige der Ermittlung des EVI. Der einzige Unterschied liegt im Ausmass der Varianz der in das lineare Loss-Integral eingehenden Normalverteilung. In den Beziehungen (3.52), (3.53) und (3.54) sind deshalb nur die $\sigma'$ durch $\sigma^*$ zu ersetzen und es gilt nach (3.51)

$$EIS(n) = \begin{cases} |t_2-t_1|\ \sigma^* L_{N^*}(D) & \text{für } m' \leq s_b \\ |t_2-t_1|\ \sigma^* L_{N^*}(-D) & \text{für } m' > s_b \end{cases} \tag{3.59}$$

mit

$$D = \frac{s_b - m'}{\sigma^*}\ .^{2} \tag{3.60}$$

Bei n=100 gilt für unser Verpackungs-Beispiel nach (3.57)

$$\sigma^{*2} = \frac{100}{225 + 100} = 0.30769$$

und

$$\sigma^* = 0.55470.$$

Nach (3.60) ist

$$D = \frac{51.04 - 51}{0.55470} = 0.0721$$

[1] Raiffa/Schlaifer, S. 297

[2] Vgl. Winkler, S. 363 f.

und nach (3.59)

$$EIS(100) = 480'000 \cdot 0.5547\ L_{N^*}(0.07)$$
$$= 266'256 \cdot 0.3649 = 97'157.$$

3.423. Binomial-Modell

3.423.1. Die Ermittlung des EVI[1]

Bei Vorliegen der Betaverteilung

$f'_\beta(\theta;r',n')$

als Apriori-Verteilung gilt für das rechte bzw. linke lineare Loss-Integral

$$L^r_\theta(s_b) = \int_{s_b}^{1} (\theta - s_b) f'_\beta(\theta;r',n')\,d\theta$$

$$= E(S')\ G_\beta(s_b;r'+1,\ n'+1) - s_b G_\beta(s_b;r',n'); \qquad (3.61)$$

$$L^l_\theta(s_b) = \int_{0}^{s_b} (s_b-\theta)\ f'_\beta(\theta;r',n')\,d\theta$$

$$= s_b F_\beta(s_b;r',n') - E(S')F_\beta(s_b;r'+1,n'+1). \qquad (3.62)$$

Mit Hilfe dieser Loss-Integrale lässt sich nun der EVI nach (3.50) wie folgt herleiten:

$$EVI = \begin{cases} |t_2-t_1|\ [E(S')G_\beta(s_b;r'+1,n'+1) - s_b G_\beta(s_b;r',n')] \text{ für } E(S') \le s_b \\ |t_2-t_1|\ [s_b F(s_b;r',n') - E'(S)F_\beta(s_b;r'+1,n'+1)] \text{ für } E(S') > s_b \end{cases} \qquad (3.63)$$

[1] Raiffa/Schlaifer, S. 263; Winkler, S. 367 ff.

Wenn die beiden Parameter der Betaverteilung $r,n \geq 1$ und ganzzahlig sind, können die Werte für $F_\beta(a;r,n)$ bzw. $G_\beta(a;r,n)$ aufgrund der Verteilungsfunktion einer binomialverteilten Zufallsvariablen bestimmt werden.

Es gilt:

$$F_\beta(a;r,n) = G_{bin}(r;a,r+n-1), \qquad \text{für } a \leq \frac{1}{2}$$

$$G_\beta(a;r,n) = G_{bin}(n;1-a,r+n-1), \qquad \text{für } a > \frac{1}{2} \quad .^1 \tag{3.64}$$

Die Bestimmung des EVI sei mittels eines Beispiels erläutert:

Bei einem Fall der Qualitätskontrolle ist

$s_b = 0.1$, $t_2 - t_1 = 300'000$

und

$f'(s) = f'_\beta(\Theta;1,15)$.

Dabei ist

$E(S') = \frac{1}{15} = 0.067$,

also

$E(S') < s_b$.

Nach (3.63) ist

$EVI = 300'000\,[0.067\, G_\beta(0.1;2,16) - 0.1\, G_\beta(0.1;1,15)]$

Nach (3.64) gilt

$F_\beta(0.1;2,16) = G_{bin}(2;0.1,17)$.

[1] Raiffa/Schlaifer, S. 217; Winkler schlägt eine andere Möglichkeit der Bestimmung von $F_\beta$ bzw. $G_\beta$ mittels der kummulierten Binomialverteilung vor, allerdings ohne einen Beweis, wie er bei Raiffa/Schlaifer zu finden ist, Winkler, S. 152 f. und S. 369.

Ferner gilt

$$G_\beta = 1 - F_\beta$$

und es ist

$$G_\beta(0.1;2,16) = 1 - G_{bin}(2;0.1,17) = 0.4818$$

und analog

$$G_\beta(0.1;1,15) = 1 - G_{bin}(1;0.1,15) = 0.2059$$

und der EVI beträgt

$$EVI = 300'000 \cdot (0.67 \cdot 0.4818 - 0.1 \cdot 0.2059)$$

$$= 3507.$$

3.423.2. Die Ermittlung des EIS[1]

Bei der Betaverteilung als Apriori-Verteilung und dem Binomial-Modell als stochastischem Modell ist die Apriori-Verteilung des Aposteriori-Erwartungswertes eine Beta-Binomialverteilung, die allgemein definiert ist als bestimmtes Integral des Produktes einer Betadichte und der Massenfunktion einer binomial-verteilten Zufallsvariablen:

$$f_{\beta bin}(r;r',n',n) = \int_0^1 f'_\beta(\theta;r',n') f_{bin}(r;\theta,n)\, d\theta \quad [2] \qquad (3.65)$$

und es gilt

$$f(E(S");r',n',n) = f_{\beta bin}(E(S")n"-r';r',n',n). \qquad (3.66)$$

$G_{\beta bin}$ bzw. $F_{\beta bin}$ können mittels folgender Relation berech-

[1] Raiffa/Schlaifer, S. 265 f.; Winkler, S. 369 f.

[2] Vgl. Raiffa/Schlaifer, S. 237

net werden:

$$p_{\beta bin}(r;r',n',n) = \frac{(r+r'-1)!(n+n'-r-r'-1)!n!(n'-1)!}{r!(r'-1)!(n-r)!(n'-r'-1)!(n+n'-1!} \quad (3.67)$$

wobei $r = 0,1,2,\ldots,$

$n = 1,2,\ldots,$

$n \geq r,\ n' > r' > 0.$ [1]

Gleichung (3.67) beruht auf einer Relation zwischen der Verteilungsfunktion der Beta-Binomialverteilung und der Verteilungsfunktion einer hypergeometrisch verteilten Zufallsvariablen.

Für das rechte bzw. linke Loss-Integral gilt

$$L^r_{E(S'')}(s_b) = E(S')\, G_{\beta bin}(s_b n''-r';r'+1,n'+1,n) - s_b\, G_{\beta bin}(s_b n''-r';r',n',n) \quad (3.68)$$

$$L^l_{E(S'')}(s_b) = s_b\, F_{\beta bin}(s_b n''-r';r',n',n) - E(S')\, F_{\beta bin}(s_b n''-r';r'+1,n'+1,n). \quad (3.69)$$

Entsprechend ist nach (3.51)

[1] Raiffa/Schlaifer, S. 237 ff.

$$EIS(n) = \begin{cases} |t_2-t_1| \cdot [E(S') G_{\beta bin}(s_b n''-r';r'+1,n'+1,n) - \\ \qquad s_b\, G_{\beta bin}(s_b n''-r';r',n',n)] , & \text{für } E(S') \le s_b , \\ |t_2-t_1| \, [s_b\, F_{\beta bin}(s_b n''-r';r',n',n) - \\ \qquad E(S')\, F_{\beta bin}(s_b n''-r';r'+1,n'+1,n)], & \text{für } E(S') > s_b . \end{cases} \tag{3.70}$$

Die Ermittlung des EIS sei mit den im vorgehenden Abschnitt verwendeten Daten für n=5 demonstriert:

Es sind $r' = 1$, $n' = 15$, $n'' = n'+n = 20$

und $s_b n''-r' = 0.1 \cdot 20 - 1 = 1.$

Nach (3.70) gilt

$$EIS(5) = 300'000\, [0.067\, G_{\beta bin}(1;2,16,5) - 0.1\, G_{\beta bin}(1;1,15,5)].$$

Ferner ist

$$G_{\beta bin}(1;r',n',n) = 1 - p_{\beta bin}(0;r',n',n).$$

Für $r' = 2$, $n' = 16$ und $n=5$ gilt nach (3.67)

$$p_{\beta bin}(0;2,16,5) = \frac{18!\ 5!\ 15!}{5!\ 13!\ 20!} = 0.5526$$

und es ist

$$G_{\beta bin}(1;2,16,5) = 1 - 0.5526 = 0.4474.$$

Analog ist

$$p_{\beta bin}(0;1.15,5) = \frac{18!\ 5!\ 14!}{5!\ 13!\ 19!} = 0.7368$$

und

$$G_{\beta bin}(1;1,15,5) = 1 - 0.7368 = 0.2632.$$

Demzufolge ist der

$$EIS\ (5) = 300'000\ [0.067 \cdot 0.4474 - 0.1 \cdot 0.2632]$$

$$= 1'097.$$

### 3.424. Poisson-Modell

#### 3.424.1. Die Ermittlung des EVI[1]

Wenn

$f'(s) = f'_\gamma(\lambda;r',t')$

ist das rechte bzw. linke Loss-Integral

$$L^r_\lambda(s_b) = \int_{s_b}^{\infty} (\lambda - s_b) f'_\gamma(\lambda;r',t')d\lambda$$

$$= E(S')\ G_{\gamma*}(s_b t';r'+1) - s\ G_{\gamma*}(s_b t';r') \qquad (3.71)$$

$$L^l_\lambda(s_b) = \int_0^{s_b} (s_b-\lambda) f'_\gamma(\lambda;r',t')d\lambda$$

$$= s_b F_{\gamma*}(s_b t';r') - E(S')F_{\gamma*}(s_b t';r'+1). \qquad (3.72)$$

Für den EVI gilt nach (3.60)

$$EVI = \begin{cases} |t_2-t_1|\ [E(S')G_{\gamma*}(s_b t';r'+1) - s_b G_{\gamma*}(s_b t';r')] & \text{für } E(S') \le s_b \\ |t_2-t_1|\ [s_b F_{\gamma*}(s_b t';r') - E(S')F_{\gamma*}(s_b t';r'+1)] & \text{für } E(S') > s_b \end{cases} \qquad (3.73)$$

[1] Raiffa/Schlaifer, S. 277 f.

Dabei gilt allgemein

$$F_{\gamma}(\lambda;r,t) = F_{\gamma^*}(\lambda t;r) \tag{3.74}$$

und

$$F_{\gamma^*}(\lambda t;r) = G_p(r;\lambda t). \tag{3.75}$$

Es sei beispielsweise

$$f'(s) = f'_{\gamma}(\lambda;2,8),$$

d.h.

$$E(S') = \frac{2}{8} = 0.25.$$

Ferner seien

$$|t_2-t_1| = 100'00$$

und $$s_b = 0.3.$$

Nach (3.73) ist der

$$EVI = 100'000\ [0.25 \cdot G_{\gamma^*}(0.3 \cdot 8;3) - 0.3 \cdot G_{\gamma^*}(0.3 \cdot 8;2)]\ .$$

Weil

$$G_{\gamma^*}(2.4;3) = 1 - F_{\gamma^*}(2.4;3)$$

und

$$F_{\gamma^*}(2.4;3) = G_p(3;2.4)$$

ist

$$G_{\gamma^*}(2.4;3) = 1 - G_p(3;2.4) = 1 - 0.4303 = 0.5697.$$

Analog ist

$$G_{\gamma^*}(2.4;2) = 1 - G_p(2;2.4) = 1 - 0.6916 = 0.3084$$

und der EVI beträgt

$$EVI = 100'000(0.25 \cdot 0.5697 - 0.3 \cdot 0.3084) = 4'991.$$

### 3.424.2. Die Ermittlung des EIS[1]

Wenn

$$f'(s) = f'_{\gamma}(\lambda;r',t')$$

und das Poisson-Modell als stochastisches Modell verwendet wird, so ist die Apriori-Verteilung des Aposteriori-Erwartungswertes eine Negativ-Binomialverteilung, die als bestimmtes Integral des Produktes einer Gammadichte mit der Massenfunktion einer poisson-verteilten Zufallsvariablen definiert ist:

$$f(E(S");r',t',t) = f_{nbin}(r^*;\frac{t}{t+t'},r'), \tag{3.76}$$

$$\text{mit} \quad r^* = s\,t"-r'. \tag{3.77}$$

Für eine Negativ-Binomialverteilung gilt

$$p_{nbin}(r;\Theta,n) = \frac{(r+n-1)!}{r!\,(n-1)!}\,\Theta^r(1-\Theta)^n \tag{3.78}$$

wobei $r = 0, 1, 2, \ldots,$

$n > 0$

$0 < \Theta < 1.$ [2]

Ferner gilt die Relation

$$G_{nbin}(r;\Theta,n) = G_{bin}(r;\Theta, r+n-1).\ ^{3} \tag{3.79}$$

Das rechte bzw. linke Loss-Integral ist

$$L^r_{E(S")}(s_b) = E(S')\,G_{nbin}(r^*;\frac{t}{t+t'},\,r'+1) - s_b G_{nbin}(r^*;\frac{t}{t+t'},\,r') \tag{3.80}$$

$$L^l_{E(S")}(s_b) = s_b F_{nbin}(r^*;\frac{t}{t+t'},r') - E(S')F_{nbin}(r^*;\frac{t}{t+t'},r'+1)\,. \tag{3.81}$$

[1] Raiffa/Schlaifer, S. 284 f.

[2] a.a.O., S. 236

[3] a.a.O.

Eingesetzt in (3.51) ergibt sich

$$EIS(t) = \begin{cases} |t_2-t_1|[E(S')G_{nbin}(r^*;\frac{t}{t+t'},r'+1)-s_b G_{nbin}(r^*;\frac{t}{t+t'},r')] \\ \qquad\qquad \text{für } E(S')\leq s_b \\ |t_2-t_1|[s_b F_{nbin}(r^*;\frac{t}{t+t'},r')-E(S')F_{nbin}(r^*;\frac{t}{t+t'},r'+1)] \\ \qquad\qquad \text{für } E(S')>s_b \end{cases}$$

(3.82)

Mit den Werten des im vorgehenden Abschnitt verwendeten Beispiels und mit t=12 gilt nach (3.77)

$r^* = 0.3 \cdot 20 - 2 = 4$

und der EIS(12) ist nach (3.82)

$$EIS(12) = 100'000[0.25\, G_{nbin}(4;\tfrac{12}{20},3) - 0.3\, G_{nbin}(4;\tfrac{12}{20},2)]$$

Nach (3.79) gilt

$G_{nbin}(4;0.6,3) = G_{bin}(4;0.6,6) = 0.5443$

bzw.

$G_{nbin}(4;0.6,2) = G_{nbin}(4;0.6,5) = 0.3370.$

Der EIS (12) beträgt

$EIS(12) = 100'000(0.25 \cdot 0.5443 - 0.3 \cdot 0.337)$

$= 3'498.$

### 3.425. Abschliessende Bemerkungen zum Konzept der konjugierten Familien von Verteilungen

Die vorgehenden Abschnitte haben gezeigt, dass das auf den ersten Blick bestechend einfach erscheinende Konzept der konjugierten Familien von Verteilungen, das eine problemlose Ermittlung der optimalen Entscheidungsregel bei ge-

gebenem Informationsumfang erlaubt und auch bei Aposteriori-Entscheidungen ohne Schwierigkeiten angewendet werden kann, im Falle der Wahl einer otpimalen Strategie sehr kompliziert werden kann. Kompliziert wird das Verfahren, sobald Erwartungswerte berechnet werden müssen. Eine Ausnahme bildet dabei das Normal-Modell, das den grossen Vorteil aufweist, dass nicht nur Apriori- und Aposteriori-Verteilung, sondern auch noch die Apriori-Verteilung des Aposteriori-Erwartungswertes der gleichen Familie von Verteilungen angehören und zudem die linearen Loss-Integrale tabelliert sind. Auch Raiffa/Schlaifer kommen zum Schluss, dass der optimale Stichprobenumfang nur im Falle des Normal-Modells (mit bekannter Varianz) mit vertretbarem Aufwand hergeleitet werden könne[1]. Sie schlagen aus diesem Grunde für das Binomial-Modell eine Approximation mit Hilfe des Normal-Modells vor, auf die hier nicht eingetreten werden soll[2].

## 4. Die Entscheidung mit zusätzlicher Informationsaktivität: variabler Informationsumfang; sequentieller Informationsentscheid

Es existiert ein Verfahren für den mehrstufigen Informationsentscheid, das eine gewisse Berühmtheit erlangt hat: die Sequentialanalyse von Wald[3]. Einen Teil ihres Bekanntheitsgrades verdankt diese Methode sicher dem Umstand, dass das Manuskript von Wald von einer amerikanischen Regierungskommission während des zweiten Weltkrieges als geheime Staatssache deklariert wurde[4], was von Anfang an

[1] Raiffa/Schlaifer, S. 131

[2] a.a.O., S. 131 ff.

[3] Vgl. Wald oder auch Mag (1); Mag (2), S. 259 ff.; Menges (3)

[4] Menges (3), S. 75

die Bedeutung des Verfahrens unterstrichen hat.

Die Sequentialanalyse ist der klassischen Statistik zuzuordnen[1], weil auch hier die Risiken von Fehlentscheidungen, α- und β-Fehler, durch einen politischen Entscheid festwerden[2]. Dabei wird der Stichprobenumfang aber nicht fixiert, sondern als Zufallsvariable behandelt. Die gesamte Anzahl der Beobachtungen ist schliesslich vom Prüfverlauf abhängig, d.h., dass beobachtet wird, wie sich die zu prüfende Grösse während des Tests ändert[3].

Aus bayesianischer Optik kann die Sequentialanalyse in zweierlei Hinsicht kritisiert werden: Erstens werden die Folgen von Fehlentscheidungen in der Analyse nicht mitberücksichtigt und zweitens keine Apriori-Wahrscheinlichkeiten verwendet.

Ein sequentielles Verfahren, das auf dem bayesianischen Standpunkt beruht, baut auf den in Abschnitt 3 dieses Teils dargestellten Methoden auf, betrachtet den Entscheid mit der Wahl der optimalen Strategie aber nicht als abgeschlossen. Nach der Durchführung der Stichprobe mit dem Umfange n* kann der Aktor die sich aufgrund der Aposteriori-Analyse ergebende Aposteriori-Verteilung als Apriori-Verteilung behandeln und ausgehend von dieser Konstellation wiederum die optimale Strategie ermitteln. Dieses Prozedere hat er solange zu wiederholen, bis n* = 0 [4].

---

[1] Vgl. S. 120

[2] Mag (1), S. 835 f.

[3] Menges (3), S. 76 f.

[4] Vgl. Winkler, S. 324

## ANHANG

### 1. Herleitung von k* für das Normal-Modell

Für die Auflösung des Ausdruckes (3)[1]

$$(3)\quad \frac{d}{dk}\left[p'(s_1)o_{21}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{k-\mu_1}{\sigma_m}} e^{-\frac{1}{2}z^2}dz + p'(s_2)o_{12}\frac{1}{\sqrt{2\pi}}\int_{\frac{k-\mu_2}{\sigma_m}}^{\infty} e^{-\frac{1}{2}z^2}dz\right] = 0$$

setzen wir

$$y_1 = \int_{-\infty}^{\frac{k-\mu_1}{\sigma_m}} e^{-\frac{1}{2}z^2}dz \quad \text{und} \quad y_2 = \int_{\frac{k-\mu_2}{\sigma_m}}^{\infty} e^{-\frac{1}{2}z^2}dz.$$

Für die Differentiation dieser Integrale, bei denen die Integrationsgrenzen von einem Parameter anhangen, gilt[2]

$$\frac{dy_1}{dk} = \int_{-\infty}^{\frac{k-\mu_1}{\sigma_m}} \frac{\delta}{\delta k} e^{-\frac{1}{2}z^2}dz + e^{-\frac{1}{2}\left[\frac{k-\mu_1}{\sigma_m}\right]^2} \frac{d\left[\frac{k-\mu_1}{\sigma_m}\right]}{dk} - e^{-\frac{1}{2}(-\infty)^2}\frac{d(-\infty)}{dk},$$

wobei

$$\int_{-\infty}^{\frac{k-\mu_1}{\sigma_m}} \frac{\delta}{\delta k} e^{-\frac{1}{2}z^2}dz = 0$$

---

[1] Vgl. S. 105

[2] Vgl. Bronstein/Semendjajew, S. 349

und $\quad e^{-\frac{1}{2}(-\infty)} \frac{d(-\infty)}{dk} = 0.$

Es ist also

$$\frac{dy_1}{dk} = e^{-\frac{1}{2}[\frac{k-\mu_1}{\sigma_m}]^2} \cdot \frac{1}{\sigma_m} .$$

Analog ist

$$\frac{dy_1}{dk} = \int_{\frac{k-\mu_2}{\sigma}}^{\infty} \frac{\delta}{\delta k} e^{-\frac{1}{2}z^2} dz + e^{-\frac{1}{2}(-\infty)\frac{d(\infty)}{dk}} - e^{-\frac{1}{2}[\frac{k-\mu_2}{\sigma_m}]^2} \cdot \frac{d[\frac{k-\mu_2}{\sigma_m}]}{dk}$$

$$= e^{-\frac{1}{2}[\frac{k-\mu_2}{\sigma_m}]^2} \cdot \frac{1}{\sigma_m} .$$

Eingesetzt in (3) ergibt sich

$$p'(s_1)o_{21}\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}[\frac{k-\mu_1}{\sigma_m}]^2} \cdot \frac{1}{\sigma_m} - p'(s_2)o_{12}\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}[\frac{k-\mu_2}{\sigma_m}]^2} \cdot \frac{1}{\sigma_m} = 0$$

oder

$$p'(s_1)o_{21}e^{-\frac{1}{2}[\frac{k-\mu_1}{\sigma_m}]^2} = p'(s_2)o_{12}e^{-\frac{1}{2}[\frac{k-\mu_2}{\sigma_m}]^2}$$

Durch Logarithmieren erhält man

$$\ln p'(s_1)+\ln o_{21}-\frac{1}{2}[\frac{k-\mu_1}{\sigma_m}]^2 \qquad \ln p'(s_2)+\ln o_{12}-\frac{1}{2}[\frac{k-\mu_2}{\sigma_m}]^2$$
$$e \qquad\qquad = e$$

oder

$$\ln p'(s_1)+\ln o_{21}-\frac{1}{2}[\frac{k-\mu_1}{\sigma}]^2 = \ln p'(s_2)+\ln o_{12}-\frac{1}{2}[\frac{k-\mu_2}{\sigma_m}]^2$$

und nach einigen Umformungen

$$k = k^* = \frac{\frac{\mu_1^2-\mu_2^2}{\sigma_m^2} - 2 \ln \frac{o_{21}p'(s_1)}{o_{12}p'(s_2)}}{\frac{2(\mu_1 - \mu_2)}{\sigma_m^2}} .$$

## 2. Herleitung von k* für das Normal-Modell als Approximation des Binomial-Modelles

Die notwendige Bedingung ist in diesem Falle erfüllt, wenn

$$(1)\left[\frac{d}{dk} o_{21}p'(s_1) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{k-\theta_1}{\sigma_1}} e^{-\frac{1}{2}z^2} dz + o_{12}p'(s_2)\frac{1}{\sqrt{2\pi}} \int_{\frac{k-\theta_2}{\sigma_2}}^{\infty} e^{-\frac{1}{2}z^2} dz\right] = 0.$$ [1]

[1] Vgl. S. 110

Im Vergleich zum im Abschnitt 1 dieses Anhanges geschilderten Verfahren ergibt sich bei der Umformung dieses Ausdruckes kein prinzipieller Unterschied:
Wenn

$$\int_{-\infty}^{\frac{k-\theta_1}{\sigma_1}} e^{-\frac{1}{2}z^2}\,dz = y_1 \quad \text{und} \quad \int_{\frac{k-\theta_2}{\sigma_2}}^{\infty} e^{-\frac{1}{2}z^2}\,dz = y_2 ,$$

kann (1) wie folgt geschrieben werden:

$$(2)\quad o_{21}p'(s_1)\,\frac{dy_1}{dk} + o_{12}p'(s_2)\,\frac{dy_2}{dk} = 0,$$

wobei

$$\frac{dy_1}{dk} = e^{-\frac{1}{2}[\frac{k-\theta_1}{\sigma_1}]^2} \cdot \frac{d[\frac{k-\theta_1}{\sigma_1}]}{dk} = e^{-\frac{1}{2}[\frac{k-\theta_1}{\sigma_1}]^2} \cdot \frac{1}{\sigma_1}$$

und

$$\frac{dy_2}{dk} = -e^{-\frac{1}{2}[\frac{k-\theta_2}{\sigma_2}]^2} \cdot \frac{d[\frac{k-\theta_2}{\sigma_2}]}{dk} = -e^{-\frac{1}{2}[\frac{k-\theta_2}{\sigma_2}]^2} \cdot \frac{1}{\sigma_2} .$$

Eingesetzt in (2):

$$o_{21}p'(s_1)\, e^{-\frac{1}{2}[\frac{k-\Theta_1}{\sigma_1}]^2} \cdot \frac{1}{\sigma_1} = o_{12}p'(s_2)\, e^{-\frac{1}{2}[\frac{k-\Theta_2}{\sigma_2}]} \cdot \frac{1}{\sigma_2} .$$

Durch Logarithmieren und nach einigen Umformungen erhält man

$$[\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}]k^2 - 2[\frac{\Theta_2}{\sigma_2^2} - \frac{\Theta_1}{\sigma_1^2}]k + \frac{\Theta_2^2}{\sigma_2^2} - \frac{\Theta_1^2}{\sigma_1^2} + 2 \ln \frac{o_{21}p'(s_1)\frac{1}{\sigma_1}}{o_{12}p'(s_2)\frac{1}{\sigma_2}} = 0.$$

Es ist also

$$k_{1,2} = \frac{2[\frac{\Theta_2}{\sigma_2^2} - \frac{\Theta_1}{\sigma_1^2}] \pm \sqrt{D}}{2[\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}]}$$

mit

$$D = 4[\frac{\Theta_2}{\sigma_2^2} - \frac{\Theta_1}{\sigma_1^2}] - 4[\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}] \; [\frac{\Theta_2^2}{\sigma_2^2} - \frac{\Theta_1^2}{\sigma_1^2} + 2 \ln \frac{o_{21}p'(s_1)\frac{1}{\sigma_1}}{o_{12}p'(s_2)\frac{1}{\sigma_2}}] .$$

Weil $\sigma_1^2 = \frac{\Theta_1(1-\Theta_1)}{n}$ und $\sigma_2^2 = \frac{\Theta_2(1-\Theta_2)}{n}$ ,

ist

$$k_{1,2} = \frac{2n\left[\frac{1}{(1-\theta_2)} - \frac{1}{(1-\theta_1)}\right] \pm \sqrt{D}}{2n\left[\frac{1}{\theta_2(1-\theta_2)} - \frac{1}{\theta_1(1-\theta_1)}\right]}$$

und

$$\sqrt{D} = 2\sqrt{n^2\left[\frac{1}{1-\theta_2} - \frac{1}{1-\theta_1}\right]^2 - n\left[\frac{1}{\theta_2(1-\theta_2)} - \frac{1}{\theta_1(1-\theta_1)}\right]\left[n\left(\frac{\theta_2}{1-\theta_2} - \frac{\theta_1}{1-\theta_1}\right) + 2\ln\frac{o_{21}p'(s_1)\sqrt{\theta_2(1-\theta_2)}}{o_{12}p'(s_2)\sqrt{\theta_1(1-\theta_1)}}\right]}.$$

Wenn man

$$a = \frac{1}{1-\theta_2} - \frac{1}{1-\theta_1} \quad \text{und} \quad b = \frac{1}{\theta_2(1-\theta_2)} - \frac{1}{\theta_1(1-\theta_1)}$$

setzt, erhält man

$$k_{1,2} = \frac{na \pm \sqrt{n^2a^2 - nb\left[na + 2\ln\frac{o_{21}p'(s_1)\sqrt{\theta_2(1-\theta_2)}}{o_{12}p'(s_2)\sqrt{\theta_1(1-\theta_1)}}\right]}}{nb} \quad . \qquad (3.31)$$

LITERATURVERZEICHNIS

ALTROGGE, Günter: Möglichkeiten und Problematik der Bewertung von (Zusatz-) Informationen mit Hilfe der Bayes-Analyse, in: ZfB 1975, S. 821 ff.

BAMBERG, Günter und COENENBERG, Gerhard: Betriebswirtschaftliche Entscheidungslehre, München 1974.

BRANT, K.E.: Derivation of Subject Probabilities to Aid in the Decision Process in Defence Weapons Acquisition, in: The Role and Effectiveness of Theories of Decision in Practice, S. 162 ff., ed. by D.J. White and K.C. Bowen, London and New York 1975.

BRONSTEIN, I.N. and SEMENDJAJEW, K.A.: Taschenbuch der Mathematik, 13. Aufl., Zürich und Frankfurt/Main 1973.

BUEHLMANN, Hans, LOEFFEL, Hans und NIEVERGELT, Erwin: Einführung in die Theorie und Praxis der Entscheidung bei Unsicherheit, 2. Aufl., Berlin, Heidelberg und New York 1969.

CARNAP, Rudolf: Statistische und induktive Wahrscheinlichkeit, in: Erkenntnisprobleme der Naturwissenschaften, S. 193 ff., hrsg. von Lorenz Krüger, Köln und Berlin 1970.

CHERNOFF, Herman and MOSES, Lincoln E.: Elementary Decision Theory, New York and London 1959.

DRUKARCZYK, Jochen: Zum Problem der Bestimmung des Wertes von Informationen, in: ZfB 1974, S. 1 ff.

DUECK, Werner und BLIEFERNICH, Manfred (Hrsg.): Operationsforschung I, Mathematische Grundlagen, Methoden und Modelle, Berlin (Ost) 1971.

EWART, Park J., FORD, James S. and LIN, Chi-Yuan: Probability for Statistical Decision Making, New Jersey 1974.

FEICHTINGER, Gustav: Zur Bayes-Analyse statistischer Entscheidungsprobleme, in: ZfB 1972, S. 449 ff.

FERSCHL, Franz: Nutzen und Entscheidungstheorie, Einführung in die Logik der Entscheidungen, Opladen 1975

GAEFGEN, Gérard: Theorie der wirtschaftlichen Entscheidung, Tübingen 1963.

GYGI, Ulrich und WEIBEL, Benedikt: Das Informationsproblem der PAN ATLANTIC Airways, in: WiST 1976, S. 194 ff.

HADLEY, G.: Introduction to Probability and Statistical Decision Theory, San Francisco usw. 1967.

HENGST, Martin: Einführung in die mathematische Statistik und ihre Anwendung, Mannheim, Wien und Zürich 1967.

JANTSCH, Erich: Technological Forecasting in Perspective, Paris 1967.

JEFFREY, Richard S.: Logik der Entscheidungen, Wien und München 1967.

KNOBLICH, Hans: Die typologische Methode in der Betriebswirtschaftslehre, in: WiST 1972, S. 141 ff.

KRAUTTER, Jochen: Marketing Entscheidungsmodelle, Wiesbaden 1973.

KREYSZIG, Erwin: Statistische Methoden und ihre Anwendungen, Nachdruck der 4. Aufl., Göttingen 1974.

KRIZ, Jürgen: Statistik in den Sozialwissenschaften, Reinbeck bei Hamburg 1973.

KUEHN, Richard: Möglichkeiten rationaler Entscheidung im Absatzsektor unter besonderer Berücksichtigung der Unsicherheit der Information, Bern und Stuttgart 1969.

LAHRMANN, Walter: Rationale unternehmerische Entscheidungen bei unvollkommener Information, Frankfurt/M. und Zürich 1973.

LIN, C.Y. and SCHICK, G.J.: Establishment of Subjective Prior Distributions with Applications in Systems Effectiveness, in: The Role and Effectiveness of Theories of Decision in Practice, S. 172 ff., ed. by D.J. White and K.C. Bowen, London and New York 1975.

LOEFFEL, Hans: Statistik und Entscheidungstheorie (Untersuchungen über das 2-Aktionen-Problem), Zürich 1970.

LUCE, Duncan R. and RAIFFA, Howard: Games and Decisions, introduction and critical survey, New York and London 1975.

MAG, Wolfgang (1): Sequentielle Informationsbeschaffung für unternehmerische Entscheidungen, in: ZfB 1973, S. 829 ff.

MAG, Wolfgang (2): Entscheidung und Information, München 1977.

MENGES, Günter (1): Grundriss der Statistik, Teil 1: Theorie, Köln und Opladen 1968.

MENGES, Günter (2): Grundmodelle wirtschaftlicher Entscheidungen, Einführung in moderne Entscheidungstheorien, Köln und Opladen 1969.

MENGES, Günter (3): Das Sequentialtestverfahren, insbesondere seine betriebswirtschaftliche Anwendung, in: ZfB 1958, S. 75 ff.

MILLER, David W. and STARR, Martin K.: Executive Decisions and Operations Research, Englewood Cliffs, N.J. 1960.

MOORE, Peter G.: Risk in business decision, London 1972.

MUELLER, Walter: Einführung in die Entscheidungstheorie, Vorlesungsskriptum, Bern 1976.

NIEVERGELT, Erwin: Ein Beitrag zur Lösung von Entscheidungsproblemen mit mehrfacher Zielsetzung, in: Die Unternehmung 1971, S. 101 ff.

OPP, Karl-Dieter: Methodologie der Sozialwissenschaften, Hamburg 1970.

PETERS, William S. and SUMMERS, George W.: Statistical Analysis for Business Decisions, Englewood Cliffs 1968.

POPPER, Karl R.: Logik der Forschung, 5. Aufl., Tübingen 1973.

RAIFFA, Howard: Decision Analysis, Introductory Lectures on Choice under Uncertainty, Menlo Park, London and Don Mills 1968.

RAIFFA, Howard and SCHLAIFER, Robert: Applied Statistical Decision Theory, paperback ed., second printing, Cambridge, Massachusetts and London 1972.

RAPPAPORT, Alfred: Sensitivity Analysis in Decision Making, in: Information for Decision Making, S. 175 ff., ed. by Alfred Rappaport, Englewood Cliffs 1975.

REICHENBACH, Hans: Die logischen Grundlagen des Wahrscheinlichkeitsbegriffs, in: Erkenntnisprobleme der Naturwissenschaften, S. 174 ff., hrsg, von Lorenz Krüger, Köln und Berlin 1970.

SACHS, Lothar: Statistische Auswertungsmethoden, 2. Aufl., Berlin, Heidelberg und New York, 1969.

SCHAEFER, Ralf, E.: Probabilistische Informationsverarbeitung, Theorie, Methoden und ein Experiment, Bern, Stuttgart und Wien 1976.

SCHLAIFER, Robert: Introduction to Statistics for Business Decisions, New York, Toronto and London 1961.

SPETZLER, Carl S. and STAËL VON HOLSTEIN, Carl-Axel S.: Probability Encoding in Decision Analysis, in: Management Science, Vol. 22, NO. 3, 1975, S. 340 ff.

STEGMUELLER, Wolfgang (1): Hauptströmungen der Gegenwartsphilosophie, Bd.1, 4. Aufl., Stuttgart 1969.

STEGMUELLER, Wolfgang (2): Rudolf Carnap: Induktive Wahrscheinlichkeit, in: Grundprobleme der grossen Philosophen, Philosophie der Gegenwart I, S. 45 ff., hrsg. von Josef Speck, Göttingen 1972.

STEGMUELLER, Wolfgang (3): Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Bd.IV: Personelle und statistische Wahrscheinlichkeit, Studienausgabe Teil A, Berlin,Heidelberg und New York 1973.

STEGMUELLER, Wolfgang (4): Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, Bd.IV: Personelle und statistische Wahrscheinlichkeit, Studienausgabe Teil D, Berlin, Heidelberg und New York 1973.

SZYPERSKI, Norbert und WINAND, Udo: Entscheidungstheorie, Stuttgart 1974.

THOMPSON, Gerald E.: Statistics for Decisions, an Elementary Introduction, Boston 1972.

TOPRITZHOFER, Edgar: Marketingentscheidungen unter Risiko - das Bayessche Konzept, in: WiST 1972, S. 301 ff. und 350 ff.

WALD, Abraham: Sequential Analysis, New York, London and Sydney 1947.

WEBER, Karl: Entscheidungsprozesse unter Verwendung des Theorems von Bayes, in: Entscheidung bei unsicheren Erwartungen, Beiträge zur Theorie der Unternehmung, S. 69 ff., hrsg. von Herbert Hax, Köln und Opladen 1963.

WEHRLI, Peter Michael: Die Lösung betriebswirtschaftlicher Entscheidungssequenzen, Diss. Bern 1976.

WEINGARTNER, Paul: Wissenschaftstheorie I, Einführung in die Hauptprobleme, Stuttgart - Bad Cannstatt 1971.

WINKLER, Robert L.: Introduction to Bayesian Inference and Decision, New York usw. 1972.

WITTGENSTEIN, Ludwig: Tractatus logico-philosophicus, Frankfurt am Main 1973.

YAMANE, Taro: Statistics, An Introducory Analysis, third ed., New York 1973.